

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 00:03:32 ; Search time 9715 Seconds  
(without alignments)  
11591.346 Million cell updates/sec

Title: US-09-436-184-3

Perfect score: 2324

Sequence: 1 cggaccgtgcaatggccag.....cttgggaactctggagaga 2324

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb.ba.\*  
2: gb.rcg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2324	100.0	2324	6	AX146795 Sequence
2	2324	100.0	2324	9	S83325 Sequence
3	2320.8	99.9	2452	6	CO871564 Sequence
4	2296.8	98.8	2449	9	HS003109 Sequence
5	1480.4	63.7	2739	4	BOVASBHY
6	1450.2	62.4	4419	10	AF289486
7	1450.2	62.4	6629	10	AF289487
8	1215.8	52.3	1970	6	CQ729666
9	947	40.7	2680	9	AF289489
10	829.6	35.7	1080	9	BC015518
11	819	35.2	2442	6	BD234762
12	819	35.2	2442	6	AR274038
13	819	35.2	2442	6	AR277619
14	819	35.2	2442	6	AR441744
15	819	35.2	2442	6	AR532049
16	819	35.2	2442	6	AX369081
17	779.4	33.5	2280	9	AF306765
18	679	29.2	2761	9	BC025236
19	455.8	19.6	2789	10	AF289488

20	412.2	17.7	2617	10	AF302653	Mus muscu
21	400.8	17.2	2770	10	BC015281	Mus muscu
22	364.4	15.7	366	6	CQ680735	Sequence
23	297	12.8	297	6	CQ731371	Sequence
24	279.2	12.0	1608	9	BC066929	Homo sapi
25	275.8	11.9	2637	10	AF302654	Mus muscu
26	267.8	11.5	524	10	AF289491	Homo sapi
27	220.2	9.5	855	9	AF224468	Mus muscu
28	218	9.4	2777	3	AF289493	Drosophil
29	215.8	9.3	566	10	AF289492	Mus muscu
30	215.8	9.3	988	10	AF223414	Mus muscu
31	215.4	9.3	597	5	EX931169	Sequence
32	215.4	9.3	652	5	CR406273	Gallus ga
33	213.6	9.2	1662	6	CQ584991	Sequence
34	208.6	9.0	3170	3	AF289494	Drosophil
35	199.2	8.6	3237	3	BT003633	Sequence
36	195.6	8.4	1014	5	EX929679	Gallus ga
37	194.2	8.4	1777	4	CFU38414	Canis fami
38	190	8.2	132868	2	AC021323	Homo sapi
39	190	8.2	158404	9	AC090094	Homo sapi
40	177.6	7.6	2432	10	AF302655	Mus muscu
41	173.8	7.5	919	10	AF221854	Mus muscu
42	173.8	7.5	1013	10	AF223413	Mus muscu
43	173.8	7.5	1733	10	AF289490	Mus muscu
44	165.2	7.1	941	9	AF224469	Mus muscu
45	165.2	7.1	1254	9	AF184241	Homo sapi

## ALIGNMENTS

RESULT 1  
AX146795  
LOCUS AX146795 2324 bp DNA linear PAT 31-MAY-2001  
DEFINITION Sequence 3 from Patent WO0135102.  
ACCESSION AX146795  
VERSION AX146795.1 GI:14285147  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Wands, J.R., de la Monte, S.M., Ince, N. and Carlson, R.I.  
TITLE Diagnosis and treatment of malignant neoplasms  
JOURNAL Patent: WO 0135102-A 3 17-MAY-2001;  
RHODE ISLAND HOSPITAL (US)  
FEATURES  
source Location/Qualifiers  
1. .2324  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

Query Match	100.0%	Score	2324	DB	6	Length	2324
Best Local Similarity	100.0%	Pred. No.	0	Mismatches	0	Indels	0
Matches	2324	Conservative	0	Gaps	0		
QY	1	CGGACCGTGCATATGGCCAGCGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCGG	60				
DB	1	CGGACCGTGCATATGGCCAGCGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCGG	60				
QY	61	GCTCCGCGACGGTAGCAGTAGTGGCGGACAGCAGCCCCCGGGGCGCGGAGAGACAA	120				
DB	61	GCTCCGCGACGGTAGCAGTAGTGGCGGACAGCAGCCCCCGGGGCGCGGAGAGACAA	120				
QY	121	AGCATGGAGGACACAGATGGAGGAAAGCGGACTCTCGGAACTTCATTCTTCAGT	180				
DB	121	AGCATGGAGGACACAGATGGAGGAAAGCGGACTCTCGGAACTTCATTCTTCAGT	180				
QY	181	GGTTTATGGTATGTCATTGCTGGCGCTGTGAGCATCTGTGCTGCTTTGTTGATC	240				
DB	181	GGTTTATGGTATGTCATTGCTGGCGCTGTGAGCATCTGTGCTGCTTTGTTGATC	240				

Qy	241	TTGTTGACTATGAGGAAGTTCTAGGAAAACTAGGAATCTATGATGCTGATGCTGATGGAG	300
Db	241	TTGTTGACTATGAGGAAGTTCTAGGAAAACTAGGAATCTATGATGCTGATGCTGATGGAG	300
Qy	301	ATTTTGATGTGATGATGCCAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTCAGAGC	360
Db	301	ATTTTGATGTGATGATGCCAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTCAGAGC	360
Qy	361	CAGCAGTCCCGCAGAGAGCGTCAGGCCACACTGAGCCCGAGGAGCAGGTTCTCTGTGG	420
Db	361	CAGCAGTCCCGCAGAGAGCGTCAGGCCACACTGAGCCCGAGGAGCAGGTTCTCTGTGG	420
Qy	421	AGGCAGAACCCCGAGAAATATCGAAGATGAACCAAAAGAACAAATTCAGTCCCTTCTCCATG	480
Db	421	AGGCAGAACCCCGAGAAATATCGAAGATGAACCAAAAGAACAAATTCAGTCCCTTCTCCATG	480
Qy	481	AAATGGTACACGCGAAGACATGTTGAGGGAGAGACTTTCGAAACGAAGAGATGGAACCAACG	540
Db	481	AAATGGTACACGCGAAGACATGTTGAGGGAGAGACTTTCGAAACGAAGAGATGGAACCAACG	540
Qy	541	GAGAACACACAGAGAGGATGATGATTTCTTATGGCGACTGATGATGATGATGATGATTTG	600
Db	541	GAGAACACACAGAGAGGATGATGATTTCTTATGGCGACTGATGATGATGATGATGATTTG	600
Qy	601	AGACCTTGGAACTGTCAATGAGAGAAACCGAGCATAGTTACCACTGGGAAGAGA	660
Db	601	AGACCTTGGAACTGTCAATGAGAGAAACCGAGCATAGTTACCACTGGGAAGAGA	660
Qy	661	CAGTTTCAAGACTGTAAATCAGGATATGGAAGAGATGATCTGTGACGAGGAATCCAG	720
Db	661	CAGTTTCAAGACTGTAAATCAGGATATGGAAGAGATGATCTGTGACGAGGAATCCAG	720
Qy	721	ATTCAGTGAACAGTAGTAGAAGATCAAGATTGCACCATGATACAGATGATGTAAAT	780
Db	721	ATTCAGTGAACAGTAGTAGAAGATCAAGATTGCACCATGATACAGATGATGTAAAT	780
Qy	781	ACCAAGTCTATGAGGAACAAGCAGTATATGAACTCTAGAAAAATGAAGGATAGAAATCA	840
Db	781	ACCAAGTCTATGAGGAACAAGCAGTATATGAACTCTAGAAAAATGAAGGATAGAAATCA	840
Qy	841	CAGAAATGCTCTCCCTGAGATTAATCCTGTAGAGATTCACAGGTAAATTTGTAGAAG	900
Db	841	CAGAAATGCTCTCCCTGAGATTAATCCTGTAGAGATTCACAGGTAAATTTGTAGAAG	900
Qy	901	AAGTAAGCATTTTCTGTGGAAGAACAGCAGGAAGTACCAACAGAAAAACAAATAGAAAAA	960
Db	901	AAGTAAGCATTTTCTGTGGAAGAACAGCAGGAAGTACCAACAGAAAAACAAATAGAAAAA	960
Qy	961	CAGATGATCCAGAACAAAGAACAAAGTTAAGAAAAAGAGCCCTAAATTTTAAATTAAT	1020
Db	961	CAGATGATCCAGAACAAAGAACAAAGTTAAGAAAAAGAGCCCTAAATTTTAAATTAAT	1020
Qy	1021	TTGATAAGCACTTTAAAGCTCAACTTGATGCTGCAGAAAAAATCCGTTAAAGGGGAAAAA	1080
Db	1021	TTGATAAGCACTTTAAAGCTCAACTTGATGCTGCAGAAAAAATCCGTTAAAGGGGAAAAA	1080
Qy	1081	TTGAGGAAGCAGTGAATGCATTTAAAGAACTPAGTACGCAAAATACCTTCAGAGTCCACGAG	1140
Db	1081	TTGAGGAAGCAGTGAATGCATTTAAAGAACTPAGTACGCAAAATACCTTCAGAGTCCACGAG	1140
Qy	1141	CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGTAATGAGG	1200
Db	1141	CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGTAATGAGG	1200
Qy	1201	TGCTACGTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCTACCTGATGCTCCCTGCAG	1260
Db	1201	TGCTACGTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCTACCTGATGCTCCCTGCAG	1260
Qy	1261	ACCTGCTGAAGCTGAGTTGAGCGTGCCTCAGACAGGCAAAATTTCTAGGTCAATGA	1320
Db	1261	ACCTGCTGAAGCTGAGTTGAGCGTGCCTCAGACAGGCAAAATTTCTAGGTCAATGA	1320

QY	1321	GAGTTCCCTGCTTACCTCGACAGATTAGTCTCAACTATTTTCCCAATGATACCTTCCCTTAA	1380
DB	1321	GAGTTCCCTGCTTACCTCGACAGATTAGTCTCAACTATTTTCCCAATGATACCTTCCCTTAA	1380
QY	1381	AAATGACCTTTGGCTGGGATACCTCTTGATAGAGATAATGACAAATGCAAGAAAGATTTT	1440
DB	1381	AAATGACCTTTGGCTGGGATACCTCTTGATAGAGATAATGACAAATGCAAGAAAGATTTT	1440
QY	1441	ATGAAGAGGTGCTGAGTGTGACACCTTAATCATGGCTTTGTCTAAAGTCCATATATGGCTTCA	1500
DB	1441	ATGAAGAGGTGCTGAGTGTGACACCTTAATCATGGCTTTGTCTAAAGTCCATATATGGCTTCA	1500
QY	1501	TCCTGAAGGCACAGAAACAAATTTGCTCAGAGCATCCCATATTTTAAAGAAAGGAATAGAAT	1560
DB	1501	TCCTGAAGGCACAGAAACAAATTTGCTCAGAGCATCCCATATTTTAAAGAAAGGAATAGAAT	1560
QY	1561	CCGGAGATCCTGGCACATGATGGGAGATTTTATTTTCCACCTGGGGATGCCATGCGA	1620
DB	1561	CCGGAGATCCTGGCACATGATGGGAGATTTTATTTTCCACCTGGGGATGCCATGCGA	1620
QY	1621	GGTTTGGGAACAAAGAGGCATATAAGTGGTATGAGCTTTGGGCAACAGAGAGGACACTTTG	1680
DB	1621	GGTTTGGGAACAAAGAGGCATATAAGTGGTATGAGCTTTGGGCAACAGAGAGGACACTTTG	1680
QY	1681	CATCTGCTGGCAACGCTCACTCTACAATGTGAATGGACTGAAAGCACAGCCTTGGTGG	1740
DB	1681	CATCTGCTGGCAACGCTCACTCTACAATGTGAATGGACTGAAAGCACAGCCTTGGTGG	1740
QY	1741	CCCCAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA	1800
DB	1741	CCCCAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA	1800
QY	1801	TCCAGATGAAGGCTTTGCACTGATGATAAAGCCAAAGGTCCTCTCTGCTCAGGATG	1860
DB	1801	TCCAGATGAAGGCTTTGCACTGATGATAAAGCCAAAGGTCCTCTCTGCTCAGGATG	1860
QY	1861	AAAACTGAGGAAAAAGGCACTGGAGCCAGTTTCACTGTGTGGACGCAAGGAAGAA	1920
DB	1861	AAAACTGAGGAAAAAGGCACTGGAGCCAGTTTCACTGTGTGGACGCAAGGAAGAA	1920
QY	1921	ATGAATTCCTGCAAGAGGACCTCTTAACCTCTACTACTAGAAAGTTTCCCGAGA	1980
DB	1921	ATGAATTCCTGCAAGAGGACCTCTTAACCTCTACTACTAGAAAGTTTCCCGAGA	1980
QY	1981	CAACAGATGACAGAGAGGACAGATCAAAATATTCATCATGCACCCCGGACTCACGTGT	2040
DB	1981	CAACAGATGACAGAGAGGACAGATCAAAATATTCATCATGCACCCCGGACTCACGTGT	2040
QY	2041	GGCGGACACAGGGCCCAAACTGCAGGCTCCGAATGCACTGGGCTTGGTGAATCCCA	2100
DB	2041	GGCGGACACAGGGCCCAAACTGCAGGCTCCGAATGCACTGGGCTTGGTGAATCCCA	2100
QY	2101	AGGAAGGCTGCAAGATTCCGATGTGCCAACGAGACAGGACCTGGGAGGAAGCAAGTGC	2160
DB	2101	AGGAAGGCTGCAAGATTCCGATGTGCCAACGAGACAGGACCTGGGAGGAAGCAAGTGC	2160
QY	2161	TCATCTTTGATGACTCTTTTGGACACGAGGATATGGCAGGATGCTCATCTTTCGGCTGA	2220
DB	2161	TCATCTTTGATGACTCTTTTGGACACGAGGATATGGCAGGATGCTCATCTTTCGGCTGA	2220
QY	2221	TATTTCATCGTGTGTGGATCGGAACTGACACCAAGAGAGAGAGAGAGAGAGAGAGAGAG	2280
DB	2221	TATTTCATCGTGTGTGGATCGGAACTGACACCAAGAGAGAGAGAGAGAGAGAGAGAGAG	2280
QY	2281	CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTCGAGAGA	2324
DB	2281	CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTCGAGAGA	2324





Qy 1381 AAAATGACCTTGGCGTGGATACCTCTTGTAGGAGATATGCAATGCAAGAAAGTTT 1440  
Db 1381 AAAATGACCTTGGCGTGGATACCTCTTGTAGGAGATATGCAATGCAAGAAAGTTT 1440  
Qy 1441 ATGAAGAGGCTGCTGAGTGTGACACCTAATGATGCTTGTGTAAGTCCATATGCTTCA 1500  
Db 1441 ATGAAGAGGCTGCTGAGTGTGACACCTAATGATGCTTGTGTAAGTCCATATGCTTCA 1500  
Qy 1501 TCCTGAAGGCACAGAAACAAATTCCTGAGAGCATCCCATATTTAAAGGAAGATAGAT 1560  
Db 1501 TCCTGAAGGCACAGAAACAAATTCCTGAGAGCATCCCATATTTAAAGGAAGATAGAT 1560  
Qy 1561 CCGGAGATCTTGGCACTGATGATGGAGATTTATTTCCACTGGGGATGCCATGCGA 1620  
Db 1561 CCGGAGATCTTGGCACTGATGATGGAGATTTATTTCCACTGGGGATGCCATGCGA 1620  
Qy 1621 GGGTTGGAAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAACAGAGGACACTTTG 1680  
Db 1621 GGGTTGGAAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAACAGAGGACACTTTG 1680  
Qy 1681 CATCTGTCTGCAACGCTCACTTACATGTGATGACTGAAGACAGCCTTGGTGA 1740  
Db 1681 CATCTGTCTGCAACGCTCACTTACATGTGATGACTGAAGACAGCCTTGGTGA 1740  
Qy 1741 CCCCCAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
Db 1741 CCCCCAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
Qy 1801 TCCGAGATGAAGGCTTGCAGTGTGATGATAAGCCAAAGGCTCTTCTGCTGAGGATG 1860  
Db 1801 TCCGAGATGAAGGCTTGCAGTGTGATGATAAGCCAAAGGCTCTTCTGCTGAGGATG 1860  
Qy 1861 AAAACCTGAGGAAAGGGAAGGAGCTGAGCCAGTTTCAAGCTGTGACAGCAAGAAAGAA 1920  
Db 1861 AAAACCTGAGGAAAGGGAAGGAGCTGAGCCAGTTTCAAGCTGTGACAGCAAGAAAGAA 1920  
Qy 1921 ATGAAATGCTGCAAGAGGAGCTCTTAAACCTGTACTTACTAGAAAGTTCCCGGAGA 1980  
Db 1921 ATGAAATGCTGCAAGAGGAGCTCTTAAACCTGTACTTACTAGAAAGTTCCCGGAGA 1980  
Qy 1981 CAACAGGATGAGAGGAGGACAGATCAATATTCATCATGCAACCCGGGACTCACGTGT 2040  
Db 1981 CAACAGGATGAGAGGAGGACAGATCAATATTCATCATGCAACCCGGGACTCACGTGT 2040  
Qy 2041 GGGCGCACACAGGGGCCCAACCTGAGGCTCCGAATGCACTGGGCTTGGTGTATCCCA 2100  
Db 2041 GGGCGCACACAGGGGCCCAACCTGAGGCTCCGAATGCACTGGGCTTGGTGTATCCCA 2100  
Qy 2101 AGGAAGGCTGCAAGATTCGATGTGCAACAGAGACAGGACCTGGGAGGAAGGCTGTC 2160  
Db 2101 AGGAAGGCTGCAAGATTCGATGTGCAACAGAGACAGGACCTGGGAGGAAGGCTGTC 2160  
Qy 2161 TCATCTTTGATGACTCTTTGAGACGAGGTATGGCAGGATGCCTCATCTTTCCGGTGA 2220  
Db 2161 TCATCTTTGATGACTCTTTGAGACGAGGTATGGCAGGATGCCTCATCTTTCCGGTGA 2220  
Qy 2221 TATTCATGCTGATGTGTGGCATCCGGAACCTGACACACAGCAGAGGAGCCTTCCAG 2280  
Db 2221 TATTCATGCTGATGTGTGGCATCCGGAACCTGACACACAGCAGAGGAGCCTTCCAG 2280  
Qy 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2324  
Db 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2324

RESULT 3  
LOCUS CQ871564  
DEFINITION Sequence 55 from Patent WO2004078783.  
ACCESSION CQ871564  
VERSION CQ871564.1 GI:52745624  
KEYWORDS linear PAT 27-SEP-2004

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Murphy, F.J., Sheehan, D.E., Keating, K.E., Hayes, I.E. and Seera, L.E.  
TITLE Enzymes involved in apoptosis  
JOURNAL Patent: WO 2004078783-A 55 16-SEP-2004;  
Biox Therapeutics Ltd (IE)  
FEATURES  
source Location/Qualifiers  
1..2452  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 99.9%; Score 2320.8; DB 6; Length 2452;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2322; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CGGACCGTGCATATGGCCAGCGTAAGAAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
Db 67 CGGACCGTGCATATGGCCAGCGTAAGAAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 126  
Qy 61 GCTCCGCGACGCTAGCAGCGTGGGCGAGCAGCGCCCGGGGCCCGGAGAGAGACAA 120  
Db 127 GCTCCGCGACGCTAGCAGCGTGGGCGAGCAGCGCCCGGGGCCCGGAGAGAGACAA 186  
Qy 121 AGCATGAGGACACAGAAATGGGAGGAAAGCGGAGCTCTCGGAACTTCACTTTACGCT 180  
Db 187 AGCATGAGGACACAGAAATGGGAGGAAAGCGGAGCTCTCGGAACTTCACTTTACGCT 246  
Qy 181 GGTATTGATGATGCTTCTGCGGCTCTGGGACATCTGTAGCTGTCTGTTGGTTTGCATC 240  
Db 247 GGTATTGATGATGCTTCTGCGGCTCTGGGACATCTGTAGCTGTCTGTTGGTTTGCATC 306  
Qy 241 TTGTTGATATGAGGAAGTTCTAGGAAATCTAGGAATCTATGATCTGTATGCTGTGATGAG 300  
Db 307 TTGTTGATATGAGGAAGTTCTAGGAAATCTAGGAATCTATGATCTGTATGCTGTGATGAG 366  
Qy 301 ATTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
Db 367 ATTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 426  
Qy 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACCTGAGCCCGGAGCAGGCTTCTGTGG 420  
Db 427 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACCTGAGCCCGGAGCAGGCTTCTGTGG 486  
Qy 421 AGGAGAAACCCAGAGATATCGAAGATGAAGCAAAAGAAATTCAGTCCCTTCTCCATG 480  
Db 487 AGGAGAAACCCAGAGATATCGAAGATGAAGCAAAAGAAATTCAGTCCCTTCTCCATG 546  
Qy 481 AAATGGTACACGAGACATGTTGAGGAGAGAGTGTGCAACAGAGAGATGAGAGAGAGAGAG 540  
Db 547 AAATGGTACACGAGACATGTTGAGGAGAGAGTGTGCAACAGAGAGATGAGAGAGAGAGAG 606  
Qy 541 GAGAACCAACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
Db 607 GAGAACCAACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 666  
Qy 601 AGACCTTGGAACTTGAAGATCTCATGAGAAACCGAGCATAGTTACCACTGGAAGAGA 660  
Db 667 AGACCTTGGAACTTGAAGATCTCATGAGAAACCGAGCATAGTTACCACTGGAAGAGA 726  
Qy 661 CAGTTTCCAGAGCTGTAATCAGGATATGAGAGAGATGATGATGATGATGATGATGATGATG 720  
Db 727 CAGTTTCCAGAGCTGTAATCAGGATATGAGAGAGATGATGATGATGATGATGATGATGATG 786  
Qy 721 ATTCAGTGAACCAAGTATGAGAGATGAAGATGCAACCATGATGATGATGATGATGATGATG 780  
Db 787 ATTCAGTGAACCAAGTATGAGAGATGAAGATGCAACCATGATGATGATGATGATGATGATG 846  
Qy 781 ACCAAGTCTATGAGGAAACAGCAGTATATGAACTCTAGAAATGAAGGGATGAGAAATCA 840

```
Db      847  ACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAAATGAAGGNTAGAAATCA 906
QY      841  CAGAAGTAACCTGCTCCCTCGAGGATAATCTCTGTAGAAAGATTCACAGGTAAATGTAGAAAG 900
Db      907  CAGAAGTAACCTGCTCCCTCGAGGATAATCTCTGTAGAAAGATTCACAGGTAAATGTAGAAAG 966
QY      901  AAGTAAGCATTTTCTCTGTGGAGAAACAGCAGGAAAGTACACACAGAAACAAATAGAAAAA 960
Db      967  AAGTAAGCATTTTCTCTGTGGAGAAACAGCAGGAAAGTACACACAGAAACAAATAGAAAAA 1026
QY      961  CAGATGATCCAGAACCAAAAGCAAAAGTTTAAAGAAAGAAAGCCTTAACCTTTTAAATTAAT 1020
Db      1027  CAGATGATCCAGAACCAAAAGCAAAAGTTTAAAGAAAGAAAGCCTTAACCTTTTAAATTAAT 1086
QY      1021  TTGATAGACTATTAAGCTGAACCTTGATGCTGCGAGAAAACCTCCGTAAAGGGGAAAAA 1080
Db      1087  TTGATAGACTATTAAGCTGAACCTTGATGCTGCGAGAAAACCTCCGTAAAGGGGAAAAA 1146
QY      1081  TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGCAAAATACCTCTCAGAGTCCAAG 1140
Db      1147  TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGCAAAATACCTCTCAGAGTCCAAG 1206
QY      1141  CAAGATATGGGAAGGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGGAGAGTAATGAGG 1200
Db      1207  CAAGATATGGGAAGGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGGAGAGTAATGAGG 1266
QY      1201  TGCTACCTGGAGCCATCGAGACCTACCAAGAGGTGGCCGCGCTACCTGATGCTCCGCGAG 1260
Db      1267  TGCTACCTGGAGCCATCGAGACCTACCAAGAGGTGGCCGCGCTACCTGATGCTCCGCGAG 1326
QY      1261  ACCTGTGAAGCTGAGTTTGAAGCGTGTGAGAGGAGCAACCAATTTCTAGGTCTATATGA 1320
Db      1327  ACCTGTGAAGCTGAGTTTGAAGCGTGTGAGAGGAGCAACCAATTTCTAGGTCTATATGA 1386
QY      1321  GAGGTTCCCTGCTTACCTCGCAGAGATTAGTTCAACTATTTTCCCAATGATATCTCTTAA 1380
Db      1387  GAGGTTCCCTGCTTACCTCGCAGAGATTAGTTCAACTATTTTCCCAATGATATCTCTTAA 1446
QY      1381  AAAATGACCTTGGCTGGATACCTCTTGTATAGAGATTAATGCAATGCAAGAAAGTTT 1440
Db      1447  AAAATGACCTTGGCTGGATACCTCTTGTATAGAGATTAATGCAATGCAAGAAAGTTT 1506
QY      1441  ATGAAGAGCTGCTCAGTGTGACACCTATGATGCTTGTCTAAAGTCCATTTATGCTTCA 1500
Db      1507  ATGAAGAGCTGCTCAGTGTGACACCTATGATGCTTGTCTAAAGTCCATTTATGCTTCA 1566
QY      1501  TCCTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCCATATTTTAAAGGAAGGAATAGAT 1560
Db      1567  TCCTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCCATATTTTAAAGGAAGGAATAGAT 1626
QY      1561  CCGGAGATCCTGGCATGTATGAGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGA 1620
Db      1627  CCGGAGATCCTGGCATGTATGAGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGA 1686
QY      1621  GGGTTGGGAACAAAGAGGCATATAGTGTATGAGCTTGGGCAACAGAGACACTTTG 1680
Db      1687  GGGTTGGGAACAAAGAGGCATATAGTGTATGAGCTTGGGCAACAGAGACACTTTG 1746
QY      1681  CATCTGTCTGGCAACCTCACTCAATATGATGAGCTGGAAGCAACAGCCTTGGTGA 1740
Db      1747  CATCTGTCTGGCAACCTCACTCAATATGATGAGCTGGAAGCAACAGCCTTGGTGA 1806
QY      1741  CCCCAGAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAAACTGGAAGTTAA 1800
Db      1807  CCCCAGAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAAACTGGAAGTTAA 1866
QY      1801  TCCGAGATGAAGCCCTTGCAAGTATGAGTAAAGCAAGCTCTCTCTGCTGAGGATG 1860
Db      1867  TCCGAGATGAAGCCCTTGCAAGTATGAGTAAAGCAAGCTCTCTCTGCTGAGGATG 1926
QY      1861  AAAACCTGAGGAAAGGGGACTGAGCCAGTTCAAGCTGTGGCAGGAAGGAAGAA 1920
```

```
Db      1927  AAAACCTGAGGGAAAAAGGGGACTTGGAGCCAGTTCAAGCTGTGCAAGCAAGGAAGAA 1986
QY      1921  ATGAAAAATGCTGCAAGAGGAGCTCTCTAAACCTGTGTACCTTACTAGAAAAAGTTCCCGAGA 1980
Db      1987  ATGAAAAATGCTGCAAGAGGAGCTCTCTAAACCTGTGTACCTTACTAGAAAAAGTTCCCGAGA 2046
QY      1981  CAACAGAGATGCAAGAGGAGACAGATCAAAATATTTCCATCATGACCCCGGACTCACGTGT 2040
Db      2047  CAACAGAGATGCAAGAGGAGACAGATCAAAATATTTCCATCATGACCCCGGACTCACGTGT 2106
QY      2041  GGGCGCACACAGGGGCCCAACAACTGCGGCTCCGAAATGCACTGGGCTTGGTATTTCCCA 2100
Db      2107  GGGCGCACACAGGGGCCCAACAACTGCGGCTCCGAAATGCACTGGGCTTGGTATTTCCCA 2166
QY      2101  AGGAAGCTGCAAGATTCGATGTGCGCAACAGAGACAGGACCTGGGAGGAAGCAAGGTGC 2160
Db      2167  AGGAAGCTGCAAGATTCGATGTGCGCAACAGAGACAGGACCTGGGAGGAAGCAAGGTGC 2226
QY      2161  TCATCTTTGATGACTCTCTTTGAGCACGAGGTATGCGAGATGCTCATCTTTCCGGCTGA 2220
Db      2227  TCATCTTTGATGACTCTCTTTGAGCACGAGGTATGCGAGATGCTCATCTTTCCGGCTGA 2286
QY      2221  TATTTCATCTGGAGTGTGGCATCCGGAACCTGACACACAGCAGAGACGAGCTTTCCAG 2280
Db      2287  TATTTCATCTGGAGTGTGGCATCCGGAACCTGACACACAGCAGAGACGAGCTTTCCAG 2346
QY      2281  CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2324
Db      2347  CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2390
```

```
RESULT 4
HSU03109
LOCUS      2449 bp mRNA linear PRI 30-NOV-1995
DEFINITION Human aspartyl beta-hydroxylase mRNA, complete cds.
ACCESSION U03109
VERSION   U03109.1 GI:458031
KEYWORDS  Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE 1 (bases 1 to 2449)
AUTHORS   Koriath, F., Gieffers, C. and Frey, J.
TITLE     Cloning and characterization of the human gene encoding aspartyl
          beta-hydroxylase
JOURNAL   Gene 150 (2), 395-399 (1994)
MEDLINE   95121937
PUBMED    7821814
REFERENCE 2 (bases 1 to 2449)
AUTHORS   Koriath, F.
TITLE     Direct Submission
JOURNAL   Submitted (03-NOV-1993) Koriath F., Fakultät fuer Chemie-Biochemie
          II, Universitaet Bielefeld, Universitaetsstrasse 25, Bielefeld,
          33615, Germany
          Location/Qualifiers
            1..2449
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="As-5"
              /cell_line="mg63"
              /cell_type="osteosarcoma"
              /clone_lib="MG63-ZAP"
            1..77
              /function="hydroxylation of aspartyl and asparaginylnyl
              residues"
              /codon_start=1
              /product="aspartyl beta-hydroxylase"
              /protein_id="AAA82108.1"
              /db_xref="GI:458032"
              /translation="MAQRKNAKSSGSSSGSGSTAGSSSPGARRETKHGHHG
```

5'UTR  
CDS

RKGLSGTSPTFTWPMVIALLGWMTSAVWVFDLVDVEVLKGLGIYDADGDGDFDVID  
 AKVLLGLKERSTSPVPEABPHTEPEQVPEAPBONIEDAEKEQIOSLHMHVH  
 AEHVSGDLQOEQDGEPOEDDEFLMATDVEDRPELPEVSHVETHSYHVEVTV  
 SQDCNQMEEHMSQENPDSSEPVVEDERLHDDTDVYQYSEQAVYELENEGIEI  
 TEVTAPPEDNPEDSQVIVIEVSIFFPVEEQEVEFPETNRKTDDEPEQAKVKKKPKLL  
 NKFDKTIKALDLAEKLRKRGKI BEAYNAFELVRYKQSPQSPRAYQACQEDDLAEKR  
 RSNELRGLATYGEVASLPLDPVADLLKLSLKRSDRQOFLGHMRGSLTLQRLVOLF  
 PNDTSLKNDLGVGLLGDNDNAKVVEEVLVTPNDGFAKVVHVGFTLKAONKIABSI  
 PYLEGESDGPDDGRFFPHLGDAMQVRGNKEAYKVELGHRGHEFASVMORSLIN  
 VNGLUKQPCPGTGYTLQVKSLSRNLWKLIRDEGLAVMDKAKGLFIPEDENLRKEDW  
 SQFTLWOOGRRNENACKGAPKTCITLLEKFPETTGRCRGQIKYISIMPHGTWHVPHGT  
 NCLRMLHGLVLPKEGCKIRCANETKTWESEKVLIFDDSPHEVWQDASSFLRIFIVD  
 VHMELTPQOORSLPFI"

3'UTR

ORIGIN

Query Match	98.8%;	Score	2296.8;	DB	9;	Length	2449;
Best Local Similarity	99.6%;	Pred. No.	0;				
Matches	2314;	Conservative	0;	Mismatches	7;	Indels	3;
				Gaps			1;
Qy	1	CGGACCGTGCATCGCCAGCGTAAGATGCAAGACAGCGGCAACAGCAGCAGCAGCG	60				
Db	67	CGGACCGTGCATCGCCAGCGTAAGATGCAAGACAGCGGCAACAGCAGCAGCAGCG	126				
Qy	61	GCTCCGGCAGCGGTAGCAGAGTGGGCGCAGCAGCGCCCGGGCCCGGAGAGACAA	120				
Db	127	GCTCCGGCAGCGGTAGCAGAGTGGGCGCAGCAGCGCCCGGGCCCGGAGAGACAA	186				
Qy	121	AGCATGGAGGACACAAGAAATGGGAGGAAAGCGGAGCTCTCGGGAACTTTCACGT	180				
Db	187	AGCATGGAGGACACAAGAAATGGGAGGAAAGCGGAGCTCTCAGGAACTTTCACGT	246				
Qy	181	GGTTATGTCATGTCATGTCGCGCTCGACATCTGTAGCTGCTGTTGGTTTCATC	240				
Db	247	GGTTATGTCATGTCATGTCGCGCTCGACATCTGTAGCTGCTGTTGGTTTCATC	306				
Qy	241	TGTTGTCATGTCAGGAACTTCTAGGAAATCTAGGAACTCTATGATGCTGATGCTGATG	300				
Db	307	TGTTGTCATGTCAGGAACTTCTAGGAAATCTAGGAACTCTATGATGCTGATGCTGATG	366				
Qy	301	ATTTTGTGTCATGTCAGGAACTTCTAGGAACTTCTAGGAACTCTATGATGCTGATGCTGATG	360				
Db	367	ATTTTGTGTCATGTCAGGAACTTCTAGGAACTTCTAGGAACTCTATGATGCTGATGCTGATG	426				
Qy	361	CAGCAGTCCCGCCAGAGAGCTGAGCCACACTGAGCCCGGAGCAGCTCTCTGTCG	420				
Db	427	CAGCAGTCCCGCCAGAGAGCTGAGCCACACTGAGCCCGGAGCAGCTCTCTGTCG	486				
Qy	421	AGGCAGAACCCAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG	480				
Db	487	AGGCAGAACCCAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG	546				
Qy	481	AAATGGTACAGCAGAAATGTTGAGGAGAGAGACTTTCGCAACAGAGATGAGCCACAG	540				
Db	547	AAATGGTACAGCAGAAATGTTGAGGAGAGAGACTTTCGCAACAGAGATGAGCCACAG	606				
Qy	541	GAGAACCAACAGAGAGATGATGATGTTCTTATGCGACCTGATGATGATGATGATG	600				
Db	607	GAGAACCAACAGAGAGATGATGATGTTCTTATGCGACCTGATGATGATGATGATG	666				
Qy	601	AGACCTCGGAACCTGAAGTATCTCATGAAGAACCGAGCATAGTTTACACSTGGAAGAGA	660				
Db	667	AGACCTCGGAACCTGAAGTATCTCATGAAGAACCGAGCATAGTTTACACSTGGAAGAGA	726				
Qy	661	CAGTTTCAAGACTGTATCAGGATATGGAAGAGATGATGTCGAGCAGGAAATCCAG	720				
Db	727	CAGTTTCAAGACTGTATCAGGATATGGAAGAGATGATGTCGAGCAGGAAATCCAG	786				
Qy	721	ATTCAGTGAACAGTGTAGAGATGAAGATTCACCATGATACAGATGATGTAACAT	780				
Db	787	ATTCAGTGAACAGTGTAGAGATGAAGATTCACCATGATACAGATGATGTAACAT	846				
Qy	781	ACCAAGTCTATGAGGAAACAGCAGTATATGAACTCTAGAAATGAAGGGATAGAAATCA	840				

Db	847	ACCAAGTCTATGAGGAAACAGCAGTATATCAACTCTAGAAATGAAGGGATAGAAATCA	906				
Qy	841	CAGAAGTAACTGCTCCCTCGAGGATAATCCCTCTAGAAGATTACACAGGTAAATGTAGAAG	900				
Db	907	CAGAAGTAACTGCTCCCTCGAGGATAATCCCTCTAGAAGATTACACAGGTAAATGTAGAAG	966				
Qy	901	AAGTAAGCATTTTTTCTGTGGAGAAACAGCAGGAGATACACACAGAAACAAATAGAAAAA	960				
Db	967	AAGTAAGCATTTTTTCTGTGGAGAAACAGCAGGAGATACACACAGAAACAAATAGAAAAA	1026				
Qy	961	CAGATGATCCAGAAACAAAAAGCAAAAGTTAAGAAAAAGAGCCCTAACTTTTAAATAAT	1020				
Db	1027	CAGATGATCCAGAAACAAAAAGCAAAAGTTAAGAAAAAGAGCCCTAACTTTTAAATAAT	1086				
Qy	1021	TTGATAGACTATTAAAGCTGAACTTGATGCTCAGAAAAAATCCCGTAAAAAGGGAAAAA	1080				
Db	1087	TTGATAGACTATTAAAGCTGAACTTGATGCTCAGAAAAAATCCCGTAAAAAGGGAAAAA	1146				
Qy	1081	TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGCGAAATACCTCTCAGAGTCCAGAG	1140				
Db	1147	TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGCGAAATACCTCTCAGAGTCCAGAG	1206				
Qy	1141	CAAGATATGGGAAGGCGCAGTGTAGGATGATTTGCTGAGAGAGAGAGAAATTAATCAGG	1200				
Db	1207	CAAGATATGGGAAGGCGCAGTGTAGGATGATTTGCTGAGAGAGAGAGAAATTAATCAGG	1266				
Qy	1201	TGCTAGCTGAGCAGCATCGAGACCTTACCAAGAGGTGGCCAGCTACCTGATGTCCTGCAG	1260				
Db	1267	TGCTAGCTGAGCAGCATCGAGACCTTACCAAGAGGTGGCCAGCTACCTGATGTCCTGCAG	1326				
Qy	1261	ACCTGCTGAAGCTGAGTTTGAAGCGTGGCTCAGACAGGCAACAAATTTCTAGGTCATATGA	1320				
Db	1327	ACCTGCTGAAGCTGAGTTTGAAGCGTGGCTCAGACAGGCAACAAATTTCTAGGTCATATGA	1386				
Qy	1321	GAGTTTCCCTGCTTACCTCAGAGATAGTTTCAACTTTTCCCATGATACCTTCCCTAA	1380				
Db	1387	GAGTTTCCCTGCTTACCTCAGAGATAGTTTCAACTTTTCCCATGATACCTTCCCTAA	1446				
Qy	1381	AAATGACCTTGGCGTGGGATACCTCTTGTAGAGATAATGACAATGCAAGAGAAAGTTT	1440				
Db	1447	AAATGACCTTGGCGTGGGATACCTCTTGTAGAGATAATGACAATGCAAGAGAAAGTTT	1506				
Qy	1441	ATGAAGAGGTGCTCAGTGTGACACCTAATGATGGCTTGTAAAGTCCATTTATGGCTTCA	1500				
Db	1507	ATGAAGAGGTGCTCAGTGTGACACCTAATGATGGCTTGTAAAGTCCATTTATGGCTTCA	1566				
Qy	1501	TCCTGAAGGACACAGAAACAAATTTGCTGAGAGCATCCCATATTTAAAGGAGAGATAGAT	1560				
Db	1567	TCCTGAAGGACACAGAAACAAATTTGCTGAGAGCATCCCATATTTAAAGGAGAGATAGAT	1626				
Qy	1561	CCGAGATCTCTGGCAGTGTAGGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGA	1620				
Db	1627	CCGAGATCTCTGGCAGTGTAGGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGA	1686				
Qy	1621	GGGTTGGGAAACAAAGAGGCAATATAAGTGTATGAGCTTGGGCAACAGAGAGGACATTTG	1680				
Db	1687	GGGTTGGGAAACAAAGAGGCAATATAAGTGTATGAGCTTGGGCAACAGAGAGGACATTTG	1746				
Qy	1681	CATCTGCTGGCAACGCTCACTCTACAATGTGATGACCTGAAAGCAGCAGCTTGTGGA	1740				
Db	1747	CATCTGCTGGCAACGCTCACTCTACAATGTGATGACCTGAAAGCAGCAGCTTGTGGA	1804				
Qy	1741	CCCCAAAGAAACCGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTAA	1800				
Db	1805	CCCCAAAGAAACCGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTAA	1863				
Qy	1801	TCCGAGATGAAGGCTTGCAGTGTATGATGAAGCCAAAGGTCTCTTCTGCTCAGGATG	1860				
Db	1864	TCCGAGATGAAGGCTTGCAGTGTATGATGAAGCCAAAGGTCTCTTCTGCTCAGGATG	1923				
Qy	1861	AAAACCTTGAGGAAAGGAGGCTGAGCCAGTTTACGCTGTGCGCAGCAAGAGAGAGAA	1920				

D	b	1924	A	NAACTGTGAGGGAAAAGGGGACTGGAGCCAGTTTCACGCCTGTGGCAGCAAGAAGAA	1983
Q	y	1921	A	TGAAATGCGTGTCAAAGAGAGCTCTAAACCTGTACTTACTTAGAAAAAGTTTCCCAGGA	1980
D	b	1984	A	TGAAATGCGTGTCAAAGAGAGCTCTAAACCTGTACTTACTTAGAAAAAGTTTCCCAGGA	2043
Q	y	1981	C	AACAGGATGCGAAGAGAGACAGATCAAAATTCCATCATGCACCCTGGGACTCACGTGT	2040
D	b	2044	C	AACAGGATGCGAAGAGAGACAGATCAAAATTCCATCATGCACCCTGGGACTCACGTGT	2103
Q	y	2041	G	CGCGCACACAGGGGCCAACCTGCAAGTCCGAATGCATCTGGGCTTGTTGATTC	2100
D	b	2104	G	CGCGCACACAGGGGCCAACCTGCAAGTCCGAATGCATCTGGGCTTGTTGATTC	2163
Q	y	2101	A	GAAGGCTGCAAGATTCGATGTCGCAACGACGACCGACCTGGGAGNAGCAAGGTGC	2160
D	b	2164	A	GAAGGCTGCAAGATTCGATGTCGCAACGACGACCGACCTGGGAGNAGCAAGGTGC	2223
Q	y	2161	T	CATCTTGTATGACTCTCTTTGAGACAAGGATATGGCAGGATCCCTCATCTTTCGGGCTGA	2220
D	b	2224	T	CATCTTGTATGACTCTCTTTGAGACAAGGATATGGCAGGATCCCTCATCTTTCGGGCTGA	2283
Q	y	2221	T	ATTCATCTGTGATGTGGCATCCGAACTGCACACAGCAGACAGCGACCTTC	2280
D	b	2284	T	ATTCATCTGTGATGTGGCATCCGAACTGCACACAGCAGACAGCGACCTTC	2343
Q	y	2281	C	AATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA	2324
D	b	2344	C	AATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA	2387
R	E	S			
B	O	V	A	S	B
L	O	C			
D	E	F	I	N	I
A	C	C	E	S	S
V	E	R	S	I	O
N	E	W	O	R	D
S	O	U	R	C	E
O	R	G	A	N	I
R	E	F	E	R	E
A	U	T	H	O	R
T	I	T	L	E	
J	O	U	R	N	A
M	E	D	L	I	N
P	U	B	L	I	S
R	E	F	E	R	E
A	U	T	H	O	R
J	O	U	R	N	A
C	O	M	M	E	N
F	E	A	T	E	S
S	O	U	R	C	E
C	O	D	S		

Db 1009 ACCGAGCTATGATGAACAA----- 1028  
QY 841 CAGAAGTAACTGCTCCCTCGAGGATTAATCTGTAGAGATTACAGAGGTAATTGTGTAAG 900  
Db 1029 -----GATCATGCTGTAGATAATTCAAACACAACTTTTGAAG 1065  
QY 901 AAGTAAGCATTTTCTCTGTGAGAGACAGCAGGAGGTACACAGAGAAACAAATAGAAAAA 960  
Db 1066 AACCGCAATGCTCCCTCGAGAGAAACAGCAAGAGTACACAGAGAAACAAATAGAAAG 1125  
QY 961 CAGATGATCCAGAACAAAGCAAAAGTTAAGAAAAAGAGCCCTAACTTTTAAATAAT 1020  
Db 1126 CAGATGAGCCAGGAAAGAAAGAAAGTGAAGAAAGAGAGCCCTAACTTTTAAATAAT 1185  
QY 1021 TTGATAAGCATTTTAAAGCTGAACTTGATGCTGAGAAAAAATCCCTGTAAGAGGAAAAA 1080  
Db 1186 TTGATAAGCATTTTAAAGCTGAGCTGATGCTGAGAAAAAATCCCTGTAAGAGGAAAAA 1245  
QY 1081 TTGAGGAAGCAGTGAATGCAATTTAAGAACTAGTACCAAAATACCTCAGAGTCCACGAG 1140  
Db 1246 TTGAGGAAGCAGTGAATGCAATTTAAGAACTAGTACCAAAATACCTCAGAGTCCACGAG 1305  
QY 1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGGAGAAATGAGG 1200  
Db 1306 CTCGGTATGGGAAGCAGTGTGAGGATGATTTGGCTGAGAGAGGAGAAATGAGG 1365  
QY 1201 TGCTAGTGGAGCCATCGAGACCTACCMAGAGGTGGCCAGGCTACCTGATGCTCCCTCAG 1260  
Db 1366 TCCTGCGCAGGCGCATCGAGACCTACCMAGAGGTGGCCAGGCTACCTGATGCTCCCTCAG 1425  
QY 1261 ACCTGCTCAAGCTGATTTGAAGCGTCTGCTCAGACAGGCAACAAATTTCTAGCTCATATGA 1320  
Db 1426 ACTTGTGAAGCTGAGCTGAGCGGAGGTGAGACAGCAGCTTTCTAGTTCACATGA 1485  
QY 1321 GAGGTTCCTCTGTTACCTCGCAGAGATTAGTTCAACTATTTCCCAATGATATCTTCTTAA 1380  
Db 1486 GAGGATCCTTGCTTACTCTCAGAAATTAGTTCAACTATTTCTCTGATGATATCTGCTTAA 1545  
QY 1381 AAATGACCTTGGCTGGGATACCTCTGTAGAGGATTAAGTGAATGCAATGCAAGAAAGTTT 1440  
Db 1546 AGAATGACCTTGGAGTTGGATACCTCTGTAGAGGATTAAGTGAATGCAATGCAAGAAAGTTT 1605  
QY 1441 ATGAAGAGTGTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTTATGGCTTCA 1500  
Db 1606 ATGAAGAGTGTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTTATGGCTTCA 1665  
QY 1501 TCCTGAAGCAGACAGAAATTTGCTGAGAGCATCCCATATTTAAGGAAGGAATAGAAAT 1560  
Db 1666 TCCTGAAGCAGACAGAAATTTGCTGAGAGCATCCCATATTTAAGGAAGGAATAGAAAT 1725  
QY 1561 CCGGAGATCTTGGCCTGATGATGGAGATTTTATTTCCACTGGGGGATGCCATGCAGA 1620  
Db 1726 CTGGGGATCCTGGTACTGATGATGGAGATTTTATTTTCCACTGGGGGATGCCATGCAGA 1785  
QY 1621 GGGTTGGGAAACAAAGGAGCATATAAGTGTATGAGCTTTGGGCACAGAGAGACACTTTG 1680  
Db 1786 GAGTTGGGAAACAAAGGAGCATATCGTGTGTACAGCTGGGCACAGAGGCGGCACACTTTG 1845  
QY 1681 CATCTGTCTGGCAACGCTCATCTACATGTGAATGGAGTGAAGACACAGCCCTTGGTGA 1740  
Db 1846 CTTCCGTGTGGCAGCGCTCACTGTACACGTGAGGATTTGAAGCCAGCGCGTGGTGA 1905  
QY 1741 CCCCAGAGAAAGCGGCTACAGAGTTAGTAAAGTCTTTAGAGAAAGAACTGGAAGTTAA 1800  
Db 1906 CCCCCAGAGAGACCGGCTACACCGAGTTAGTGAATGCTTTAGAGAAAGAACTGGAAGTTAA 1965  
QY 1801 TCCGAGATGAAGCCCTTGCAGTGAATGATGAATAAGCCAAAGGTCTTCTTCTGCTGAGGATG 1860  
Db 1966 TCAGAGATGAAGCCCTTGCAGCAATGATGAATAAGCCAAAGGTCTTCTTCTGCTGAGGATG 2025  
QY 1861 AAAACCTGAGGAAAAAGGGGAGTGAAGCCAGTTACGCTGTGGCAGGCAAGGAGAGAA 1920

2026 AGAACCTGAGGAAAAAGGATGAGTGAATTCAGCTGTGGCAGCAAGGAAGAAAA 2085  
QY 1921 ATGAAATGCTGCAAGAGAGCTCTCTAAACCTGTACTTACTAGAAAAAGTTTCCCGAGA 1980  
Db 2086 ATGAAATGCTGTAAGAGAGCTCTCTAAAGCTGTTCTTACTAGATAAATTCCTGAGA 2145  
QY 1981 CAACAGGATGCAAGAGAGAGAGATCAATATTTCCATCATGACCCCGGACTCAGGTGT 2040  
Db 2146 CAACAGGATGCGAAGAGAGAGAGATCAATATTTCCATCATGACCCCGGAAACCCACGTGT 2205  
QY 2041 GCGCCGACACAGGCGCCACCACTGAGCTCCGAATGACCTGGCTTGGTGTATTCCTCA 2100  
Db 2206 GCGCGCACACAGGAGCCACGAACTGTAGCTCCGAATGACCTGGCTTGGTGTATTCCTCA 2265  
QY 2101 AGGAAGGCTGCAAGATTTGATGTGCAACAGAGACAGGACCTGGAGAGGAGCAAGGTGC 2160  
Db 2266 AGGAAGGCTGCAAGATTCGATCGGCAACAGAGACAGGACCTGGAGAGGAGCAAGGTGC 2325  
QY 2161 TCATCTTGTGATGACTCTTTGAGCAGGAGTATGCGAGATGCGCTCATCTTTCGGCTCA 2220  
Db 2326 TCATCTTGTGATGACTCTTTGAAACAGGAGTGTGGCAGGATGCGCGTCTTTCGGCTCA 2385  
QY 2221 TATTTCATCTGATGATGTGTCGATCCGGAATCGACACAGCAGAGACAGCAGCTTCCAG 2280  
Db 2386 TCTTCATTTGATGATGTGTCGATCCGACCTGAGCTGACACCCCATCAGAGACGAGCTTCCCTG 2445  
QY 2281 CAATTTAGCATGAATTCATGCAAGTTGGGAACTCTGAGAGAGA 2324  
Db 2446 CAATTTAGCAT---TCTGTGCGGCTCAGGACACGCTGGAGAGA 2486

RESULT 6  
AF289486 4419 bp mRNA linear ROD 18-DEC-2000  
LOCUS Mus musculus aspartyl beta-hydroxylase 4.5 kb transcript (Asph)  
DEFINITION mRNA, complete cds, alternatively spliced.  
ACCESSION AF289486  
VERSION AF289486.1 GI:11878109  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 4419)  
AUTHORS Dinchuk,J.E., Henderson,N.L., Burn,T.C., Huber,R., Ho,S.P., Link,J., O'Neil,K.T., Focht,R.J., Scully,M.S., Hollis,J.M., Hollis,G.F. and Friedman,P.A.  
TITLE Aspartyl beta-hydroxylase (Asph) and an evolutionarily conserved isoform of Asph missing the catalytic domain share exons with junctin  
J. Biol. Chem. 275 (50), 39543-39554 (2000)  
JOURNAL MEDLINE 20564328  
PUBMED 10956665  
REFERENCE 2 (bases 1 to 4419)  
AUTHORS Dinchuk,J.E., Henderson,N.L., Burn,T.C., Hollis,G.F. and Friedman,P.A.  
TITLE Direct Submission  
JOURNAL Submitted (26-JUL-2000) Applied Biotechnology, DuPont Pharmaceuticals, P.O. Box 80400, Wilmington, DE 19880-0400, USA  
FEATURES  
source Location/Qualifiers  
1..4419  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="BALB/c"  
/db\_xref="taxon:10090"  
/chromosome="8"  
/map="8q12 at STS WT-11767"  
/tissue\_type="liver"  
1..4419  
/gene="Asph"  
1..174  
/gene="Asph"  
175..2394







Qy	1743	CCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTTAAAGAAACCTGGAGTTAATC	1802
Db	1849	CCACAGGAGACTGGCTACACAGAGCTAGTGAAGTCTTTTAAAGAAACCTGGAGTTAATC	1908
Qy	1803	CGAGATGAAGCCCTTGAGTGATGATGAATGAAGCCAAAGTCTCTCTGCTGCTGAGGATGAA	1862
Db	1909	CGTGATGAAGCCCTTGAGTGATGATGAATGAAGCCAAAGTCTCTCTGCTGCTGAGGAGAA	1968
Qy	1863	AACCTGAGGGAAGAAAGGGGACTGGAGCCAGTTTCAAGCTGCGGAGCAAGAAAGAAAT	1922
Db	1969	AACCTTGGGAGAGGCGGACTGGAGCCAGTTTCAAGCTGCGGAGCAAGAAAGAAAT	2028
Qy	1923	GAAATGCTGCAAGAGAGCTCTAAACCTGTACCTTACTAGAAAGTTCCCGGAGACA	1982
Db	2029	GAGAATGCTGTAAGAGAGCGCTTAAGACCTGTGCTTTACTAGAAAGTTTCCGAAACA	2088
Qy	1983	ACAGGATGCAAGAGAGACAGATCAATATATTCATCATGACACCCCGGACTCATCGTGTGG	2042
Db	2089	ACAGGATGCAAGAGAGACAGATCAATATATTCATCATGACACCCCGGACTCATCGTGTGG	2148
Qy	2043	CCGCACACAGGGCCCAACAACTGAGGCTCGAATGACCTGGGCTTGGTGTATCCCAAG	2102
Db	2149	CCGCATACAGGACCCACAACTGAGGCTCGAATGACCTGGGCTTGGTGTATCCCAAG	2208
Qy	2103	GAGGCTCGAGATTCGATGTCGCAAGAGACAGGACCTGGGAGGAGGCAAGTGTCTC	2162
Db	2209	GAGGCTCGAGATTCGATGTCGCAAGAGACAGGACCTGGGAGGAGGCAAGTGTCTC	2268
Qy	2163	ATCTTTGATGACTCTTTGAGCAGCAGGATGGCAGGATGCTCATCTTTCGGCTGATA	2222
Db	2269	ATCTTTGATGACTCTTTGAGCAGCAGGATGGCAGGATGCTCATCTTTCGGCTGATA	2328
Qy	2223	TTTCATGCTGGATGTGGTGCATCCGAACTGACACACAGAGAGCAGCCTTCCAGCA	2282
Db	2329	TTTCATGCTGGATGTGGTGCATCCGAACTGACACACAGAGAGCAGCCTTCCGCA	2388
Qy	2283	ATTAGCATGATTCATGCAAGCTTGG	2309
Db	2389	ATTGAAAGGCACTGACGAGACTTGG	2415
RESULT 7			
AF289487			
LOCUS			
DEFINITION	AF289487 6629 bp mRNA linear ROD 18-DEC-2000		
ACCESSION	Mus musculus aspartyl beta-hydroxylase 6.6 kb transcript (Asph)		
VERSION	AF289487		
KEYWORDS	mRNA, complete cds; alternatively spliced.		
SOURCE	AF289487.1 GI:11878111		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Mus musculus		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 6629)		
JOURNAL	Dinchuk, J.E., Henderson, N.L., Burn, T.C., Huber, R., Ho, S.P., Link, J., O'Neil, K.T., Focht, R.J., Scully, M.S., Hollis, J.M., Hollis, G.F. and Friedman, P.A.		
MEDLINE	Aspartyl beta-hydroxylase (Asph) and an evolutionarily conserved isoform of Asph missing the catalytic domain share exons with junction		
PUBMED	J. Biol. Chem. 275 (50), 39543-39554 (2000)		
REFERENCE	20564328		
AUTHORS	10956665		
TITLE	Dinchuk, J.E., Henderson, N.L., Burn, T.C., Hollis, G.F. and Friedman, P.A.		
JOURNAL	Direct Submission		
MEDLINE	Submitted (26-JUL-2000) Applied Biotechnology, DuPont Pharmaceuticals, P.O. Box 80400, Wilmington, DE 19880-0400, USA		
PUBMED	1-6629		
FEATURES	Location/Qualifiers		
source	1-6629		
organism	Mus musculus		
mol_type	"mRNA"		

gene	/strain="BALB/c"	
5'UTR	/db_xref="taxon:10090"	
	/chromosome="8"	
	/map="8q12 at STS WT-11767"	
	/tissue_type="liver"	
	1..6629	
	/gene="Asph"	
	1..174	
	/gene="Asph"	
	175..2394	
	/gene="Asph"	
	/note="BAH; asparaginyl beta-hydroxylase; Asph; alternatively spliced"	
	/codon_start=1	
	/product="aspartyl beta-hydroxylase 6.6 kb transcript"	
	/protein_id="AAG40809.1"	
	/db_xref="GI:11878112"	
	/translation="MAPRKAAGGGSSSGSGSGSGSGSPSTGSSSGSSPCARRE AKGHGKNGRRGGISGGSFTWPMVITALLGWTVSVVWFDLVYDEVILGKLVYDAD GDGDFDADAKVLLGLKERSPSERTFPPEAEATHAELEQAPEGADIQNVDEVEKQIO SLLQESVHTDHLDEADGLAGPQPEVEDFLTVTSDSDRPFDELDPEDEPVEEIEDTYHVE FTSQNHNDMEEMTEQNSDPSEAVTDAGVLLPHABEVHRQDYDEVPSPSEHGV RKIENAVNAPEELVRKYQSPRARVKAOCEDDLAEKORNEVLRRAIETVQEAADL DPATDLKLSLKRSSRQQLGHRGSLTLQRLVOLFPSDTTLKNDLGVGLLGD NDASKVVEEVNLNTPNDGFAKYHGFILKAQNKLSIESIPYLKESIESDPTDGRF YFLGDAMORVGNKEAYKWEIGHKRGHFASVWQKSLNVNGLKAQPMWTPRETQYTE LKSLERLNWKLIRDEGLAMDKAGLFLPEDENREKGDWDSOFTLWQGRKNENACKG APTKALLKPFSETTCGRGQIKYSIMHPGTHVWPTGTNCRMLHLGLVLPKRGCK IRCANETRTWEEGKVLIPDDSPHEVWQDASSFRLLFIVDVWHPBLTPQRRSLPAI"	
	2395..6629	
3'UTR	/gene="Asph"	
	4379..4384	
	/gene="Asph"	
	6573..6578	
	/gene="Asph"	
	polyA_signal	
	polyA_signal	
ORIGIN		
	Query Match	62.4%; Score 1450.2; DB 10; Length 6629;
	Best Local Similarity	79.5%; Pred. No. 1.1e-303;
	Matches 1833; Conservative	0; Mismatches 378; Indels 96; Gaps 6;
QY	15	GCCAGCGTAAAGATCCAAAGACAGCGGCAACAGCAGCAGCAGCGGCTCCGCGACGGT 74
DB	193	GCCAAAGGCGCGCGCGGCAACAGCAGCAGCAGCGGCGGCGGCGGCGGCTCGGCGACGGT 252
QY	75	AGCACGAGTGTGGGCGACGACG-----AGCCCCGGGCCCCGAGAGACAAAG 122
DB	253	AGCCGAGCAGCGGTAGCAGCAGTAGCCCGGGGCTCGGAGAGAGGCAAG 312
QY	123	CATGAGGACACAGATGGGAGGAAAGCGGACTCTCGGGAACCTTCATCTTCAGGTGG 182
DB	313	CACGAGGACACAGATGGGAGGAGAGAGGAGATTCAGAGGGTCTCTTTTCACATGG 372
QY	183	TTTATGTGATGTCATGCTGGGCGCTCGGACATCTGTAGCTGTGCTTGGTTGATCTT 242
DB	373	TTTATGTGATGTCATGCTGGGCGCTCGGACATCTGTGCTGTGCTGTTGATCTT 432
QY	243	GTTCACATGAGGAAGTTCAGGAAACTAGGAACTATGATGCTGATGCTGATGATGAGAT 302
DB	433	GTTCATGATGAGGAAGTTCAGGAAACTAGGAGTCTATGATGCGGATGCGGATGAGAC 492
QY	303	TTTGTGATGATGATGCCAAAGTTTATTAGGACTTTAAAGAGAGAGATCTACTTCAGAGCCA 362
DB	493	TTTGTGATGATGATGCCAAAGTTTATTAGGCTTTAAGAAAGATCTCTCTCTGAGAGG 552
QY	363	GCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGAGGTTCTCTGTGGAG 422
DB	553	ACATTTCCACC---GGAGGCAGAGACTCAGCTGAGCTGGAGGAGGAGGAGGAGGAGGAG 609
QY	423	GCAGAACCCGAGAAATATCGAAGATGAGCAAAAGAAACAAATTCAGTCCCTTCTCCATGAA 482
DB	610	GCAGACATCCAGAAATGTTGAAGATGAAGTAAGGAACAAATTCAGTCCCTTCTTCAGGAA 669

QY 483 ATGTTACACCGGAAACATGTTGAGGGAGAGACTTGGCAACAAGAGATGGAACCCACAGA 542  
Db 670 TCAGTACACACAGACCATG-----ACCTTGAGCAGATGGCGTGGCGGA 714  
QY 543 GAAACACACAGAGAGATGATGATTTCTTATGCGGACTGATGATGATGATGATGATGAG 602  
Db 715 GAAACACAGCGGAGGTTGAGGACTTCTCTACAGTGACCGCAGTGTGATGATGATGATGAG 774  
QY 603 ACCCTGGAACCTGAAATGATCTCATGAAGAAACCGAGCAGTGTACACACGTGGAAGAGACA 662  
Db 775 GACCTGGAACCCGGGACAGTTTATGAAGAAATCGAGGATACCTTACCATGTGGAAGACACA 834  
QY 663 GTTTCACAGACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAGAT 722  
Db 835 GCATCGCAGAACCATCCAAATGACATGGAGATGACGATGAGCAGGAAATTCAGAT 894  
QY 723 TCCAGTGAACAGTGTAGAGATGAAGATGACCATGATGATGATGATGATGATGATGATGAT 782  
Db 895 CCAAGTGAAGCAGTGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 954  
QY 783 CAAGTGTATGAGGAAACAGCAGTATATGAACCTCTAGAAATGAAAGGATGAGAAATCACA 842  
Db 955 CAAGACTATGATGAAC---CAGTATATGAACCTTACAGGACAGGAGGAGGTCGAGATTTCA 1011  
QY 843 GAAGTAATCTCTCCCTCGAGGATATCTGTAGAAAGATTCACAGGTAAATTTAGAGAA 902  
Db 1012 -----GATAACACCATAGATGATTCAGCATATCTCAGAGAA 1050  
QY 903 GTAAGCATTTTCTCTGTAAGNACAGCAGGAGATGATCCACCNGAACAATAGAAACA 962  
Db 1051 ATAAATGTCCTCTCTGTGAGNACAGCAGACACACACC----- 1091  
QY 963 GATGATCCAGAACAAAGAAAGTTAAAGAAAGAGGCTTAACTTTTAAATAATTT 1022  
Db 1092 -----AGTTAAGAAAGAGAGGCTTAACTTTTGAACAAATTT 1128  
QY 1023 GATAAGACTATTAAAGCTGAACCTTGATGCTGCAGAAATACTCCGTAAAAGGGGAAAAAT 1082  
Db 1129 GATAAGAACTATTAAAGCTGAGTGTGATGCTGCAGAAAGCTCCGGAAAAGGGGTAATTT 1188  
QY 1083 GAGGAACAGTGAATGCAATTTAAAGAACTAGTAGCAAAATACCTCAGAGTCCACGAGA 1142  
Db 1189 GAGGAACAGTGAATGCAATTTGAGAACTGTTGCAANGTACCTCAGAGGCCACGAGA 1248  
QY 1143 AGATATGGGAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGGAGAGATGATGAGTG 1202  
Db 1249 AGATATGGCAAGCGCAGTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1308  
QY 1203 CTAGTGGAGCCATCGAGACCTTACCAAGAGTGCCAGCCTTACTGATGCTCCCTGACAG 1262  
Db 1309 CTGGCAGGGCCATCGAGACCTTACCAGGAGGACCGCCAGCCTGCTGATGCCCTACAGAC 1368  
QY 1263 CTGCTGAAGCTGAGTTGAGGCTCGCTCAGACAGGCAACAATTTCTAGTTCATATGAGA 1322  
Db 1369 CTGCTGAAGCTGAGTTGAGGCTGAGGAGGCTGGAACGCGCAGGCTTCTAGTTCATATGAGA 1428  
QY 1323 GGTTCCTGCTTACCTGCTGAGAGATTTAGTTCAACTATTTCCCAATGATCTTCTTAAAA 1382  
Db 1429 GGTTCCTTACTTACCTTCCAGAGACTAGTTCAGTGTTCCTAGTGTACTACTTAA 1488  
QY 1383 AATGACCTTGGCTGGGATACCTCTTGATAGGAGATGATGATGATGATGATGATGATGAT 1442  
Db 1489 AACGACCTTGGCTGGGATACCTCTTGATGGGAGACACGACAGTGCAGGAGGATTTAC 1548  
QY 1443 GAAGAGTGTGAGTGTGACACCTAATGATGCTTTGCTTAAAGTCCATATGCTTCTATC 1502  
Db 1549 GAAGAGTGTAAATGTGACACCAATGATGGCTTCGCTTAAAGTGTGATACGGCTTCTATC 1608  
QY 1503 CTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCATATTTAAAGGAAAGAAATAGATCC 1562  
Db 1609 CTGAAGGCACAGAAACAAATTTGCTGAGAGCATTCCTTAAAGGAAAGAAATCGAATCT 1668

QY 1563 CGAGATCTGGCACTGATGATGGAGATTTTATTTCCACCTGGGGGATGCCATGCCAGAG 1622  
Db 1669 GGGGACCTTGGCAGGATGATGCGGTTTACTTTCCACTTGGGGGATGCCATGCCAGAGG 1728  
QY 1623 GTTGGGAAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAACAAGAGAGACACTTTGCA 1682  
Db 1729 GTCGGGAAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAACAAGAGAGACACTTTGCC 1788  
QY 1683 TCTGTCTGGCAACGCTCACTCTACAATGTGAATGAGTGAAGACACAGCCTTGGTGGACC 1742  
Db 1789 TCTGTCTGGCAGCGTTCCCTCTACAATGTGAATGAGTGAAGGCTCAGCGCTGGTGACA 1848  
QY 1743 CCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAATC 1802  
Db 1849 CCCAGGAGACTGGCTACACAGAGCTAGTGAAGTCTTTAGAGAGAACTGGAAGTTAATC 1908  
QY 1803 CGAGATGAGGCTTGGCTGATGATGATGAAGCAAGGCTCTTCTCTGCTGAGGATGAA 1862  
Db 1909 CGTATGAGGCTTCTATGATGATGAAGCAAGGCTCTTCTCTGCTGAGGATGAA 1968  
QY 1863 AACCTGAGGAAACGGGCTGAGGAGCTGAGCCAGTTTCACTGTGGCAGCAAGGAAAGAAAT 1922  
Db 1969 AACCTTGGGAGAGGCGGCTGAGCCAGTTTCACTGTGGCAGCAAGGAAAGAAAT 2028  
QY 1923 GAAATGCTTGCAGAGGAGCTCTAAACCTGTACTTACTAGAAAAGTTTCCCGGAGACA 1982  
Db 2029 GAGAATGCTTGAAGAGGCTTAAAGGCTTAAAGCTTGTCTTACTAGAAAAGTTTCCGAAACA 2088  
QY 1983 ACAGATGACAGAGGAGCAGATCAAAATTTCCATCATGACCCCGGCTCAGCTGTGG 2042  
Db 2089 ACAGATGACAGAGGAGCAGATCAAAATTTCCATCATGACCCCGGCTCAGCTGTGG 2148  
QY 2043 CCGCACACAGGCGCCACAACTGCGAGGCTCCGAATGCACTTGGGCTTGGTGTATCCCAAG 2102  
Db 2149 CCGCATACAGGACCCCAAACTGCGAGGCTCCGAATGCACTTGGGCTTGGTGTATCCCAAG 2208  
QY 2103 GAAGGCTGCAAGATTCGATGTGTCACACGAGACCGAGGACCTGGGAGGAAAGCAAGTGTCT 2162  
Db 2209 GAAGGCTGCAAGATTCGATGTGTCACACGAGACCGAGGACCTGGGAGGAAAGCAAGTGTCT 2268  
QY 2163 ATCTTTGATGACTCTTTTGGAGCAGGATGATGGAGGATGCTTCTTTCGGCTGATA 2222  
Db 2269 ATCTTTGATGACTCTTTTGGAGCAGGATGATGGAGGATGCTTCTTTCGGCTGATA 2328  
QY 2223 TTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2282  
Db 2329 TTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2388  
QY 2283 ATTTAGCATGAATTCATGCAAGCTTGG 2309  
Db 2389 ATTTGAAAGGCACTGACGAGACTTGG 2415

## RESULT 8

LOCUS CQ729666 1970 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 15600 from Patent WO02068579.  
ACCESSION CQ729666  
VERSION CQ729666.1 GI:42301377  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
TITLE Kits, such as nucleic acid arrays, comprising a majority of  
humanexons or transcripts, for detecting expression and other uses  
thereof  
JOURNAL Patent: WO 02068579-A 15600 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES  
Location/Qualifiers  
1..1970

ORIGIN		/organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"	
Query Match		52.3%; Score 1215.8; DB 6; Length 1970;	
Best Local Similarity		90.2%; Pred. No. 7.5e-253;	
Matches 1299; Conservative		0; Mismatches 141; Indels 0; Gaps 0;	
Qy	849	ACTGCTCCCTCAGGAGTAACTCTGTAGAGATTTCACAGGTAAATGTAGAGAGTAAGC	908
Db	531	ACTGAGCCGAGGAGCAGGTTCTGTGGAGCGAACCAGAAATATCGAAGATGAAGCA	590
Qy	909	ATTTTCTCTGTGAAGAACAGCAGAAAGTACCACAGAAACAATAAGAAAAACAGATGAT	968
Db	591	AAAGAACAAATTCAGTCCCTTCTCCATGAAATGGTACACGACAGAACATGTTGAGGAGAA	650
Qy	969	CCAGAAACAAAGCAAAAGTTAAGAAAAGAGCCTAAACTTTTAAATAATTTGATAAG	1028
Db	651	GACTTGCACACAGAAAGATGGACCCACAGAGGAACCAACAAAGAGGATGATGATTTCTT	710
Qy	1029	ACTATTAAAGCTGAATTTGATGCTGCAGAAAACCTCCGTTAAAGGGGAAAAATTTGAGGAA	1088
Db	711	ATGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	770
Qy	1089	GCAGTGAATGCATTTAAAGAACTAGTAGCGCAAAATPACCTTCAGAGTCCACGAGCAAGATAT	1148
Db	771	GCAGTGAATGCATTTAAAGAACTAGTAGCGCAAAATPACCTTCAGAGTCCACGAGCAAGATAT	830
Qy	1149	GGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAGTAATGAGGTGCTACGT	1208
Db	831	GGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAGTAATGAGGTGCTACGT	890
Qy	1209	GGACCCATCGAGACCTACCAAGAGTGCAGAGCTTACTGATGTCCTGCAGACCTCTCTG	1268
Db	891	GGACCCATCGAGACCTACCAAGAGTGCAGAGCTTACTGATGTCCTGCAGACCTCTCTG	950
Qy	1269	AAGCTGATGTTGAAGCGTCGTCCAGACGAGCAACAATTTCTAGTGCATATGAGAGGTTCC	1328
Db	951	AAGCTGATGTTGAAGCGTCGTCCAGACGAGCAACAATTTCTAGTGCATATGAGAGGTTCC	1010
Qy	1329	CTGCTTACCTGCGAGAGATGATTTCAACTATTTCCCAATGATGATGATGATGATGATGATGAT	1388
Db	1011	CTGCTTACCTGCGAGAGATGATTTCAACTATTTCCCAATGATGATGATGATGATGATGATGAT	1070
Qy	1389	CTTGCGTGGGATACCTCTTCATAGGAGATTAATCAATGCAAGAAAGTTTATGAAGAG	1448
Db	1071	CTTGCGTGGGATACCTCTTCATAGGAGATTAATCAATGCAAGAAAGTTTATGAAGAG	1130
Qy	1449	GTGCTGATGTGACACCTAATGATGGCTTTGCTAAAGTCCATTAATGCTTCATCCTGAAG	1508
Db	1131	GTGCTGATGTGACACCTAATGATGGCTTTGCTAAAGTCCATTAATGCTTCATCCTGAAG	1190
Qy	1509	GCACAGAACAAATTTGCTGAGAGATCCCATATTTAAAGAGAGAGTAATCGGAGAT	1568
Db	1191	GCACAGAACAAATTTGCTGAGAGATCCCATATTTAAAGAGAGAGTAATCGGAGAT	1250
Qy	1569	CCTGCGCTGATGATGGAGATTTTATTTCCACCTGGGGATGCGATGAGAGGTTGGG	1628
Db	1251	CCTGCGCTGATGATGGAGATTTTATTTCCACCTGGGGATGCGATGAGAGGTTGGG	1310
Qy	1629	AACAAAGGCGATATAAGTGGTATGAGCTTGGGCAACAGAGAGGACATTTGCAATCTGTC	1688
Db	1311	AACAAAGGCGATATAAGTGGTATGAGCTTGGGCAACAGAGAGGACATTTGCAATCTGTC	1370
Qy	1689	TGGCAACGCTCACTCTACAAATGTGAATGGACTGAAGCAGACGCTTGGTGCACCCCAAA	1748
Db	1371	TGGCAACGCTCACTCTACAAATGTGAATGGACTGAAGCAGACGCTTGGTGCACCCCAAA	1430
Qy	1749	GAAAGGCGCTACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAATCCGAGAT	1808
Db	1431	GAAAGGCGCTACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAATCCGAGAT	1490
Qy	1809	GAAGGCTTGCAGTGATGATGAATAAGCAAAAGGTCTCTTCTGCTGAGGATGAACACCTG	1868
Db	1491	GAAGGCTTGCAGTGATGATGAATAAGCAAAAGGTCTCTTCTGCTGAGGATGAACACCTG	1550
Qy	1869	AGGAAAAAGGGGACTGGAGCCAGTTTCAGCTCTGCGAGCAAGAAAGAAAGTGAATAAT	1928
Db	1551	AGGAAAAAGGGGACTGGAGCCAGTTTCAGCTCTGCGAGCAAGAAAGTGAATAATGAATAAT	1610
Qy	1929	GCCTGCAAAAGGAGCTCTTAAACCTGTACTAGAAAGTTCCCGGAGACAACAGGA	1988
Db	1611	GCCTGCAAAAGGAGCTCTTAAACCTGTACTAGAAAGTTCCCGGAGACAACAGGA	1670
Qy	1989	TGCAAGAGGAGACAGATCAATAATTTCCATCATGCAACCCGGGACTCACGTGTGGCCGAC	2048
Db	1671	TGCAAGAGGAGACAGATCAATAATTTCCATCATGCAACCCGGGACTCACGTGTGGCCGAC	1730
Qy	2049	ACAGGCCCCCAAACTGCGAGCTCCGAAATGCACTGGGCTTGGTGTATTTCCCAAGGAAGGC	2108
Db	1731	ACAGGCCCCCAAACTGCGAGCTCCGAAATGCACTGGGCTTGGTGTATTTCCCAAGGAAGGC	1790
Qy	2109	TGCAAGATTCGATGTGTCACAGACGAGCAAGCACTGGGAGGAAGCAAGGTGCTCATCTTT	2168
Db	1791	TGCAAGATTCGATGTGTCACAGACGAGCAAGCACTGGGAGGAAGCAAGGTGCTCATCTTT	1850
Qy	2169	GATGACTCTTTGAGCAGCAGGATATGCGAGGATGCTCATCTTTCCGCTGATATTCATC	2228
Db	1851	GATGACTCTTTGAGCAGCAGGATATGCGAGGATGCTCATCTTTCCGCTGATATTCATC	1910
Qy	2229	GTGAGTGTGGCATCCGAACTGACACACAGACGAGCGAGCCCTCCAGCAATTTAG	2288
Db	1911	GTGAGTGTGGCATCCGAACTGACACACAGACGAGCGAGCCCTCCAGCAATTTAG	1970
RESULT 9			
AF289489 2680 bp mRNA linear PRI 18-DEC-2000			
LOCUS			
DEFINITION Homo sapiens aspartyl beta-hydroxylase 2.8 kb transcript mRNA,			
complete cds; alternatively spliced.			
ACCESSION AF289489			
VERSION AF289489.1 GI:11878115			
KEYWORDS			
SOURCE Homo sapiens (human)			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1 (bases 1 to 2680)			
Dinchuk, J.E., Henderson, N.L., Burn, T.C., Huber, R., Ho, S.P.,			
Link, J., O'Neill, K.T., Focht, R.J., Scully, M.S., Hollis, J.M.,			
Hollis, G.P. and Friedman, P.A.			
AUTHORS			
Aspartyl beta -hydroxylase (Asph) and an evolutionarily conserved			
isoform of Asph missing the catalytic domain share exons with			
junction			
J. Biol. Chem. 275 (50), 39543-39554 (2000)			
JOURNAL			
MEDLINE 20564328			
FUBMED 10956665			
REFERENCE			
2 (bases 1 to 2680)			
Dinchuk, J.E., Henderson, N.L., Burn, T.C., Hollis, G.P. and			
Friedman, P.A.			
AUTHORS			
Direct Submission			
TITLE			
Submitted (26-JUL-2000) Applied Biotechnology, DuPont			
JOURNAL			
Pharmaceuticals, P.O. Box 80400, Wilmington, DE 19880-0400, USA			
FEATURES			
Location/Qualifiers			
1..2680			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/chromosome="8"			
/map="8q12 at STS WT-11767"			
/tissue_type="liver"			
1..11			
12..953			
5'UTR			
CDS			
note="Humbig; asparaginyl beta-hydroxylase; Asph;			
non-catalytic isoform; alternatively spliced"			



Db	718	GAGAACCAACAAACAGAGGATGATGAGTTTCTTATGGCGACTGATGTAGATCATGATTTG	777
QY	601	AGACCTTGGAACTCTGAAGTATCTCATGAAGAAACCGAGCATAGTTTACAAGTGAAGAGA	660
Db	778	AGACCTTGGAACTCTGAAGTATCTCATGAAGAAACCGAGCATAGTTTACAAGTGAAGAGA	837
QY	661	CAGTTTCAACAAGACTGTAAATCAGGATATGAAGAGATGATGCTGAGCAGGAAAATCCAG	720
Db	838	-----CAG 840	
QY	721	ATTCCAGTGAACCTGAAGTAGTAGAAGATGAAAGATTGCACCATGATACAGATGATGTAAACAT	780
Db	841	ATTCCAGTGAACCTGAAGTAGTAGAAGATGAAAGATTGCACCATGATACAGATGATGTAAACAT	900
QY	781	ACCAAGTCTATGAGGAACCAAGCAGTATATGAACCTCTAGAAAAATGAAGGGATAGAAATCA	840
Db	901	ACCAAGTCTATGAGGAACCAAGCAGTATATGAACCTCTAGAAAAATGAAGGGATAGAAATCA	960
QY	841	CAGAAGTAAGTCTCCCTCCCTGAGGATTAATCCTGTAGAAGATTACACAGTAAATTTGTAGNAG	900
Db	961	CAGAAGTAAGTCTCCCTCCCTGAGGATTAATCCTGTAGAAGATTACACAGTAAATTTGTAGNAG	1020
QY	901	AAGTAAAGCATTTTCTGTGTGGAAGAACAGCAGGAAAGTACCACCAGAAACAAATAGAAAAA	960
Db	1021	AAGTAAAGCATTTTCTGTGTGGAAGAACAGCAGGAAAGTACCACCAGAAACAAATAGAAAAA	1080
RESULT	11		
LOCUS	BD234762	2442 bp	DNA linear PAT 17-JUL-2003
DEFINITION	Gene encoding cancer marker labyrinthine.		
ACCESSION	BD234762		
VERSION	BD234762.1	GI:33044532	
KEYWORDS	JP 2002512005-A/1.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 2442)		
TITLE	Radosevich,J.A.		
JOURNAL	Gene encoding cancer marker labyrinthine		
COMMENT	Patent: JP 2002512005-A 1 23-APR-2002;		
	JAMES A RADOSEVICH		
	OS Homo sapiens (human)		
	PN JP 2002512005-A/1		
	PD 23-APR-2002		
	PF 11-MAR-1999 JP 2000536866		
	PR 17-MAR-1998 US 09/040485		
	PI JAMES A RADOSEVICH		
	PC C12N15/09,C07K14/47,C07K16/18,C12P21/08,C12Q1/02,C12Q1/68// PO (C12P21/08,C12R1/91,C12N15/00		
	CC Gene encoding cancer marker labyrinthine		
	FH Key Location/Qualifiers		
	FT CDS (70)..(834).		
FEATURES	Location/Qualifiers		
source	1..2442		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	35.2%	Score 819;	DB 6; Length 2442;
Best Local Similarity	98.8%;	Pred. No. 8.4e-167;	
Matches	825; Conservative	0; Mismatches 10; Indels 0; Gaps 0	
QY	121	AGCATGAGACACACAAGTATGGGAGGAAAGCGGAGCTCTCGGAACTTCATCTTCACGT	180
Db	5	AGCTTGAAGACACACAAGTATGGGAGGAAAGCGGAGCTCTCAGGAACCTCATCTTCACGT	64
QY	181	GGTTATGGTATTGCAATTCCTGGCGCTCTGGACATCTGTAGCTGTCTGTTGGTTGATC	240
Db	65	GGTTATGGTATTGCAATTCCTGGCGCTCTGGACATCTGTAGCTGTCTGTTGGTTGATC	124





AR441744	GI:42667904	US 6509448-A 1791 21-JAN-2003;	US 6667154-A 1791 23-DEC-2003;
Version	Unknown.	Location/Qualifiers	Location/Qualifiers
Keywords	Unknown.	1. .2442	1. .2442
Source	Unknown.	/organism="unknown"	/organism="unknown"
Organism	Unknown.	/mol_type="genomic DNA"	/mol_type="genomic DNA"
Reference	Unclassified.		
Authors	Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S., Carter, D., Retter, W., Mannion, J. and Pan, L.		
Title	Compositions and methods for the therapy and diagnosis of lung cancer		
Journal	Patent: US 6667154-A		
Features	Patent: US 6667154-A 1791 23-DEC-2003;		
Source	Location/Qualifiers		
	1. .2442		
Origin	Query Match		
	Best Local Similarity 98.8%; Pred. No. 8.4e-167;		
	Matches 825; Conservative 0; Mismatches 10; Indels 0; Gaps 0;		
QY	121 AGCATGAGGACACAAAGATGGAGAAAGCGGACTCTCGGAACTTCATTCTTACGT	180	121 AGCATGAGGACACAAAGATGGAGAAAGCGGACTCTCGGAACTTCATTCTTACGT
DB	5 AGCTTGAAGGACACAAAGATGGAGAAAGCGGACTCTCAGGAACCTTCATTCTTACGT	64	5 AGCTTGAAGGACACAAAGATGGAGAAAGCGGACTCTCAGGAACCTTCATTCTTACGT
QY	181 GGTTTATGCTGATTCGATTCGCTGGCGTCTGGACATCTGTAGCTGTCTGTTGGTTC	240	181 GGTTTATGCTGATTCGATTCGCTGGCGTCTGGACATCTGTAGCTGTCTGTTGGTTC
QY	65 GGTTTATGCTGATTCGATTCGCTGGCGTCTGGACATCTGTAGCTGTCTGTTGGTTC	124	65 GGTTTATGCTGATTCGATTCGCTGGCGTCTGGACATCTGTAGCTGTCTGTTGGTTC
DB	241 TTGTTGACTATGAGAAAGTCTAGGAAAACCTAGGAACTCTATGATGCTGATGATGATG	300	241 TTGTTGACTATGAGAAAGTCTAGGAAAACCTAGGAACTCTATGATGCTGATGATGATG
QY	125 TTGTTGACTATGAGAAAGTCTAGGAAAACCTAGGAACTCTATGATGCTGATGATGATG	184	125 TTGTTGACTATGAGAAAGTCTAGGAAAACCTAGGAACTCTATGATGCTGATGATGATG
QY	301 ATTTTATGCTGATTCGATTCGCTGGCGTCTGGACATCTGTAGCTGTCTGTTGGTTC	360	301 ATTTTATGCTGATTCGATTCGCTGGCGTCTGGACATCTGTAGCTGTCTGTTGGTTC
DB	185 ATTTTATGCTGATTCGATTCGCTGGCGTCTGGACATCTGTAGCTGTCTGTTGGTTC	244	185 ATTTTATGCTGATTCGATTCGCTGGCGTCTGGACATCTGTAGCTGTCTGTTGGTTC
QY	361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGGTTCTCTGG	420	361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGGTTCTCTGG
DB	245 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGGTTCTCTGG	304	245 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGGTTCTCTGG
QY	421 AGGACACACACAGAGAGGATTCGAGAACTAGGAACTATGATCTGATGATGATGATG	480	421 AGGACACACACAGAGAGGATTCGAGAACTAGGAACTATGATCTGATGATGATGATG
DB	305 AGGACACACACAGAGAGGATTCGAGAACTAGGAACTATGATCTGATGATGATGATG	364	305 AGGACACACACAGAGAGGATTCGAGAACTAGGAACTATGATCTGATGATGATGATG
QY	481 AAATGTTACACGACGACATGTTGAGGAGAGAGACTTGCACAGAGAGATGGACCCAC	540	481 AAATGTTACACGACGACATGTTGAGGAGAGAGACTTGCACAGAGAGATGGACCCAC
DB	365 AAATGTTACACGACGACATGTTGAGGAGAGAGACTTGCACAGAGAGATGGACCCAC	424	365 AAATGTTACACGACGACATGTTGAGGAGAGAGACTTGCACAGAGAGATGGACCCAC
QY	541 GAGAACACACAGAGAGGATTCGAGAACTAGGAACTATGATCTGATGATGATGATG	600	541 GAGAACACACAGAGAGGATTCGAGAACTAGGAACTATGATCTGATGATGATGATG
DB	425 GAGAACACACAGAGAGGATTCGAGAACTAGGAACTATGATCTGATGATGATGATG	484	425 GAGAACACACAGAGAGGATTCGAGAACTAGGAACTATGATCTGATGATGATGATG
QY	601 AGACCTGGAACCTGGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTTGAAGA	660	601 AGACCTGGAACCTGGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTTGAAGA
DB	485 AGACCTGGAACCTGGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTTGAAGA	544	485 AGACCTGGAACCTGGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTTGAAGA
QY	661 CAGTTTCAACAGACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG	720	661 CAGTTTCAACAGACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG
DB	545 CAGTTTCAACAGACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG	604	545 CAGTTTCAACAGACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG
QY	721 ATTCAGTGAACAGTAGTAGAGAGTGAAGATTCGACCATGATACAGATGATGTAACAT	780	721 ATTCAGTGAACAGTAGTAGAGAGTGAAGATTCGACCATGATACAGATGATGTAACAT
DB	605 ATTCAGTGAACAGTAGTAGAGAGTGAAGATTCGACCATGATACAGATGATGTAACAT	664	605 ATTCAGTGAACAGTAGTAGAGAGTGAAGATTCGACCATGATACAGATGATGTAACAT
QY	781 ACCAAGTCTATGAGGAAACAGAGTATATGAACCTCTAGAAATGAAGGATGAGAAATCA	840	781 ACCAAGTCTATGAGGAAACAGAGTATATGAACCTCTAGAAATGAAGGATGAGAAATCA
DB	665 ACCAAGTCTATGAGGAAACAGAGTATATGAACCTCTAGAAATGAAGGATGAGAAATCA	724	665 ACCAAGTCTATGAGGAAACAGAGTATATGAACCTCTAGAAATGAAGGATGAGAAATCA
QY	841 CAGAAGTAACTGCTCCCTGAGGATTAATCTGTAGAGATTCACAGGTAATTTGTAAG	900	841 CAGAAGTAACTGCTCCCTGAGGATTAATCTGTAGAGATTCACAGGTAATTTGTAAG
DB	725 CAGAAGTAACTGCTCCCTGAGGATTAATCTGTAGAGATTCACAGGTAATTTGTAAG	784	725 CAGAAGTAACTGCTCCCTGAGGATTAATCTGTAGAGATTCACAGGTAATTTGTAAG
QY	901 AAGTAAGCATTTTCTGTGGAAGACAGCAGAGAGTACCACAGAAACAAATAG	955	901 AAGTAAGCATTTTCTGTGGAAGACAGCAGAGAGTACCACAGAAACAAATAG
DB	785 AAGTAAGCATTTTCTGTGGAAGACAGCAGAGAGTACCACAGAAACAAATAG	839	785 AAGTAAGCATTTTCTGTGGAAGACAGCAGAGAGTACCACAGAAACAAATAG
RESULT 14			
AR441744			
LOCUS			
DEFINITION			
Sequence 1791 from patent US 6667154.			
2442 bp DNA linear PAT 20-FEB-2004			



```
Qy 901 AAGTAAGCATTTTCTCTGGAAGACACGACGAGTATACCCAGAACAAATAG 955
Db 785 AAGTAAGCATTTTCTCTGGAAGACACGACGAGTATACCCAGTACTTAAG 839

RESULT 15
AR532049
LOCUS AR532049
DEFINITION Sequence 1 from patent US 6727080..
ACCESSION AR532049
VERSION AR532049.1 GI:53920638
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2442)
AUTHORS Radosevich J
TITLE Gene encoding labyrinthin, a marker for cancer
JOURNAL Patent: US 6727080-A 1 27-APR-2004;
FEATURES
    Location/Qualifiers
        source
            1..2442
                /organism="unknown"
                /mol_type="genomic DNA"

ORIGIN
Query Match 35.2%; Score 819; DB 6; Length 2442;
Best Local Similarity 98.8%; Pred. No. 8.4e-167;
Matches 825; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 121 AGCATGGAGGACACAAATGCGAGGAGGCGGACTCTCGGAACTTCATCTTCACGT 180
Db 5 ACCTTGAAGACACAAATGCGAGGAGGCGGACTCTCAGGAACCTTCATCTTCACGT 64

Qy 181 GGTATTATGGTGAATTCGATCTGGGGCTCTGGACATCTGTAGCTGTGCTTTGTTGATC 240
Db 65 GGTATTATGGTGAATTCGATCTGGGGCTCTGGACATCTGTAGCTGTGCTTTGTTGATC 124

Qy 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATCTGTGATGGTGGAG 300
Db 125 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATCTGTGATGGTGGAG 184

Qy 301 ATTTTGATGTGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 360
Db 185 ATTTTGATGTGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 244

Qy 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTGTGG 420
Db 245 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTGTGG 304

Qy 421 AGGAGAACCCCGAGATATCGAAGATGAAGCAAGAAAGCAAAATTCAGTCCCTTCCATG 480
Db 305 AGGAGAACCCCGAGATATCGAAGATGAAGCAAGAAAGCAAAATTCAGTCCCTTCCATG 364

Qy 481 AAATGGTACACGACAAACATGTTGAGGAGAGAGACTTGCACAAAGAGATGGACCCACAG 540
Db 365 AAATGGTACACGACAAACATGTTGAGGAGAGAGACTTGCACAAAGAGATGGACCCACAG 424

Qy 541 GAGAACCAACCAAGAGGATGATGATTTCTTATGGGACTGTAGTGTAGATGATGATTTG 600
Db 425 GAGAACCAACCAAGAGGATGATGATTTCTTATGGGACTGTAGTGTAGATGATGATTTG 484

Qy 601 AGACCTGGAACTGAAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGGAGAGA 660
Db 485 AGACCTGGAACTGAAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGGAGAGA 544

Qy 661 CAGTTTCAAGAATCTGAATCAGGATATGGAAGAGATGATCTCTGAGCAGGAAATCCAG 720
Db 545 CAGTTTCAAGAATCTGAATCAGGATATGGAAGAGATGATCTCTGAGCAGGAAATCCAG 604

Qy 721 ATTCAGTGAACAGTAGAAGATGAAGATTGCACCATGATACAGATGATGATTAACAT 780
Db 605 ATTCAGTGAACAGTAGAAGATGAAGATTGCACCATGATACAGATGATGATTAACAT 664
```

```
Qy 781 ACCAAGTCTATGAGGAAACAAGCAGTATATGAAACCTCTAGAAAATGAAGGGATAGAATCA 840
Db 665 ACCAAGTCTATGAGGAAACAAGCAGTATATGAAACCTCTAGAAAATGAAGGGATAGAATCA 724

Qy 841 CAGAAGTAACCTGCTCCCTGAGGATAATCCTGTAGAAGATTCACAGTAATTTGTAGAAG 900
Db 725 CAGAAGTAACCTGCTCCCTGAGGATAATCCTGTAGAAGATTCACAGTAATTTGTAGAAG 784

Qy 901 AAGTAAGCATTTTCTCTGGAAGACACGACGAGTATACCCAGAACAAATAG 955
Db 785 AAGTAAGCATTTTCTCTGGAAGACACGACGAGTATACCCAGAACAAATAG 839

Search completed: March 25, 2005, 12:32:14
Job time : 9726 secs
```

10-10-2010 2:10 PM

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 07:40:08 ; Search time 6777 Seconds  
(without alignments)  
13053.174 Million cell updates/sec

Title: US-09-436-184-3  
Perfect score: 2324  
Sequence: 1 cggaccgtgcaatggccag.....cttgggaactctggagaga 2324

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_est3.\*  
4: gb\_est4.\*  
5: gb\_est5.\*  
6: gb\_est6.\*  
7: gb\_est7.\*  
8: gb\_est8.\*  
9: gb\_est9.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2054.4	88.4	2126	9	AY405267 Homo sapi
2	1497.6	64.4	2023	9	AY405268 Pan trogl
3	1450	62.4	4406	3	AK030293 Mus muscu
4	1374.4	59.1	3924	3	AK049506 Mus muscu
5	1332.6	57.3	2854	3	AK035735 Mus muscu
6	1264	54.4	2153	9	AY405269 Mus muscu
7	905.2	39.0	1050	4	BM806632 AGENCOURT
8	873.4	37.6	1011	5	BM806632 AGENCOURT
9	834.6	35.9	1085	5	BM806632 AGENCOURT
10	787.6	33.9	859	1	BU130952 AGENCOURT
11	719.8	31.0	930	5	BU130952 AGENCOURT
12	697	30.0	782	7	CK724966 UI-H-E01
13	689	29.5	689	4	BM790814 K-EST0070
14	667.8	28.7	918	5	BU166117 AGENCOURT
15	657.6	28.3	765	6	CD643948 AGENCOURT
16	639	27.5	737	7	BU166117 AGENCOURT
17	628.4	27.0	1014	5	BU190749 AGENCOURT
18	627.4	27.0	729	7	BU190749 AGENCOURT
19	627.2	27.0	959	5	BU190749 AGENCOURT
20	626.4	27.0	868	5	BU190749 AGENCOURT
21	621.6	26.7	719	7	CR630295 DKF2p4691
22	614.4	26.4	651	2	AW996828 QV3-BN004
23	612	26.3	750	6	CA317663 UI-M-FW0
24	601.4	25.9	603	5	BP336604 BP336604

25	595.4	25.6	721	4	BI333062	BI333062 602996326
26	595	25.6	799	4	BG917932	BG917932 602820860
27	593.8	25.6	876	7	CO733503	CO733503 SUT02005
28	581.4	25.0	767	1	AU132666	AU132666 AU132666
29	580.4	25.0	582	5	BP270196	BP270196 BP270196
30	578.4	24.9	581	5	BP266690	BP266690 BP266690
31	572	24.6	583	5	BP364957	BP364957 BP364957
32	571.4	24.6	770	7	CN368823	CN368823 328775548
33	567	24.4	589	7	CN258117	CN258117 170005318
34	566.2	24.4	583	5	BP328261	BP328261 BP328261
35	562.8	24.2	1100	5	BQ278505	BQ278505 AGENCOURT
36	546	23.5	554	1	AL700393	AL700393 DKF2p686K
37	542.4	23.3	711	7	CN368832	CN368832 170004182
38	534.6	23.0	715	4	BG701542	BG701542 602682786
39	530.4	22.8	629	7	CN483754	CN483754 hW36h12.1
40	523.4	22.5	589	4	BG720220	BG720220 602692270
41	522.6	22.5	537	2	BE935834	BE935834 QV2-NN004
42	521.6	22.4	553	4	BI035330	BI035330 QV2-NN200
43	519.4	22.3	811	5	BU144875	BU144875 603228482
44	517	22.2	517	5	BX479249	BX479249 DKF2p686L
45	514.6	22.1	1048	2	BE876059	BE876059 601485742

## ALIGNMENTS

RESULT 1  
AY405267  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
FEATURES  
source  
gene  
ORIGIN

AY405267 2126 bp DNA linear GSS 16-DEC-2003  
Homo sapiens ASPH gene, VIRTUAL TRANSCRIPT, partial sequence,  
Genomic survey sequence.  
AY405267 GI:39761241  
GSS.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2126)  
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,  
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sainsky, J.J.,  
Adams, M.D. and Cargill, M.  
Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
Science 302 (5652), 1960-1963 (2003)  
14671302  
2 (bases 1 to 2126)  
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,  
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sainsky, J.J.,  
Adams, M.D. and Cargill, M.  
Direct Submission  
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
This sequence was made by sequencing genomic exons and ordering  
them based on alignment.  
Location/Qualifiers  
1..2126  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
<1..>2126  
/gene="ASPH"  
/locus\_tag="HWM2167"

Query Match 88.4%; Score 2054.4; DB 9; Length 2126;  
Best Local Similarity 96.7%; Pred. No. 0;  
Matches 2055; Conservative 0; Mismatches 70; Indels 0; Gaps 0;  
QY 12 ATGGCCACGCTTAAAGTCCAGAGCGGCGGACGAGCGGCTCCGGCAGC 71  
|||||

Db 1 ATGCCCCAGCTTAAGATGCCAAGAGCAGCGCAACAGCAGCAGCAGCGGCTCCGCGAGC 60  
Qy 72 GGTAGCA CGAGTGGGCGACAGCAGCGCCCGGGGCCCGGAGAGAGACAAAGCATGAGGA 131  
Db 61 GGTAGCA CGAGTGGGCGACAGCAGCGCCCGGGGCCCGGAGAGAGACAAAGCATGAGGA 120  
Qy 132 CACAAGATGGGGAAGGCGACTCTCGGGAACTTCTTTCAGCTGTTTATGGTG 191  
Db 121 CACAAGATGGGGAAGGCGGACTCTCAGGAACCTTCACTTTCAGCTGTTTATGGTG 180  
Qy 192 ATTCAATTCCTGGCGCTGGACATCTGTAGCTGTGTTTGGTTGATCTTGTGACTAT 251  
Db 181 ATTCAATTCCTGGCGCTGGACATCTGTAGCTGTGTTTGGTTGATCTTGTGACTAT 240  
Qy 252 GAGGAAGTTCTAGGAATACTAGGAATCTATGATGCTGATGGTGATGGAGATTTTGATG 311  
Db 241 GAGGAAGTTCTAGNN 300  
Qy 312 GATGATGCCAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCCG 371  
Db 301 NNN 360  
Qy 372 CCAGAAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTGTGGAGGCGAAGACC 431  
Db 361 CCAGAAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTGTGGAGGCGAAGACC 420  
Qy 432 CAGAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATGAATGGTACAC 491  
Db 421 CAGAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATGAATGGTACAC 480  
Qy 492 GCAGAACATGTTGAGGGAAGAGACTTGCAACAAGATGACCCACAGGAGAACCCACAA 551  
Db 481 GCAGAACATGTTGAGGGAAGAGACTTGCAACAAGATGACCCACAGGAGAACCCACAA 540  
Qy 552 CAAGAGGATGATGATGTTTATGGCGATGATGTAGATGATGATTTGAGACCCCTGGAA 611  
Db 541 CAAGAGGATGATGATGTTTATGGCGATGATGTAGATGATGATTTGAGACCCCTGGAA 600  
Qy 612 CTTGAAGTATCTCATGAAGAACCGAGCATAGTTTACCACGTTGGAAGAGACAGTTTCACAA 671  
Db 601 CTTGAAGTATCTCATGAAGAACCGAGCATAGTTTACCACGTTGGAAGAGACAGTTTCACAA 660  
Qy 672 GACTCTAATCAGGATATGGAAGAGATGATGTCTGAGCAGGAGAAATCCAGATTCAGTGAA 731  
Db 661 GACTCTAATCAGGATATGGAAGAGATGATGTCTGAGCAGGAGAAATCCAGATTCAGTGAA 720  
Qy 732 CCAGTAGTAGAAGATGAAGATTGACCATGATACAGATGATGTAAACATCCAGTCTAT 791  
Db 721 CCAGTAGTAGAAGATGAAGATTGACCATGATACAGATGATGTAAACATCCAGTCTAT 780  
Qy 792 GAGGAACAGCAGTATGAACCTCTAGAAATGAAGGATAGAAATCACAGAAAGTAACT 851  
Db 781 GAGGAACAGCAGTATGAACCTCTAGAAATGAAGGATAGAAATCACAGAAAGTAACT 840  
Qy 852 GCTCCCTCTGAGGATAAATCTGTAGAAGATTCACAGGTAATTTGTAGAAGAGTAAAGCAT 911  
Db 841 GCTCCCTCTGAGGATAAATCTGTAGAAGATTCACAGGTAATTTGTAGAAGAGTAAAGCAT 900  
Qy 912 TTTCTGTGGAGAGACAGCAGGAAGTACACAGAAACAAATAGAAACACAGATGATCCA 971  
Db 901 TTTCTGTGGAGAGACAGCAGGAAGTACACAGAAACAAATAGAAACACAGATGATCCA 960  
Qy 972 GAACAAAAGCAAAAGTTTAAAGAAAAGAGCCCTTAAATTTTAAATTTTAAATTTTAAAGACT 1031  
Db 961 GAACAAAAGCAAAAGTTTAAAGAAAAGAGCCCTTAAATTTTAAATTTTAAATTTTAAAGACT 1020  
Qy 1032 ATTTAAAGCTGAATCTGATGCTGCAGAAAAAATCCGTTAAAGGGGAAAAAATTTGAGGAAGCA 1091  
Db 1021 ATTTAAAGCTGAATCTGATGCTGCAGAAAAAATCCGTTAAAGGGGAAAAAATTTGAGGAAGCA 1080  
Qy 1092 GTGAATGCAATTTAAAGAACTAGTACCAATACCTTCAGGTCACAGCAGCAAGATATGGG 1151  
Db 1081 GTGAATGCAATTTAAAGAACTAGTACCAATACCTTCAGGTCACAGCAGCAAGATATGGG 1140

RESULT 2  
AY405268  
LOCUS

AY405268 2023 bp DNA linear GSS 16-DEC-2003

Qy 1152 AAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGGAGAGAGTAATGAGGTGCTACGTGA 1211  
Db 1141 AAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGGAGAGAGTAATGAGGTGCTACGTGA 1200  
Qy 1212 GCCATCGAGACCTTACCAAGAGGTGGCCAGCTTACCTGATGTCCCTGCAGACCTGCTGAAG 1271  
Db 1201 GCCATCGAGACCTTACCAAGAGGTGGCCAGCTTACCTGATGTCCCTGCAGACCTGCTGAAG 1260  
Qy 1272 CTGAGTTTCAAGCGTGCCTCAGACAGGCAACAATTTCTAGGTATATGAGAGGTTCCCTG 1331  
Db 1261 CTGAGTTTCAAGCGTGCCTCAGACAGGCAACAATTTCTAGGTATATGAGAGGTTCCCTG 1320  
Qy 1332 CTTTACCTCGAGAGATTTAGTTCAACTATTTCCTTCCCAATGATCTTCTTAAAAAATGACCTT 1391  
Db 1321 CTTTACCTCGAGAGATTTAGTTCAACTATTTCCTTCCCAATGATCTTCTTAAAAAATGACCTT 1380  
Qy 1392 GGGCTGGGATACCTCTTGTAGTAGAGATATGCAATGCAAGAAAGATTTATGAAGAGTG 1451  
Db 1381 GGGCTGGGATACCTCTTGTAGTAGAGATATGCAATGCAAGAAAGATTTATGAAGAGTG 1440  
Qy 1452 CTGAGTGTGACACCTAATGATGCTTTGCTAAAGTCCATTATGGCTTCATCTCTGAAGCA 1511  
Db 1441 CTGAGTGTGACACCTAATGATGCTTTGCTAAAGTCCATTATGGCTTCATCTCTGAAGCA 1500  
Qy 1512 CAGAACAAAAATGCTCGAGAGCATCCCATATTTAAAGGAAAGGAATAGAATCCGGAGATCCT 1571  
Db 1501 CAGAACAAAAATGCTCGAGAGCATCCCATATTTAAAGGAAAGGAATAGAATCCGGAGATCCT 1560  
Qy 1572 GGCACGTATGATGGGAGATTTATTTCCACCTGGGGATGCCATGCAGAGGTTTGGAAAC 1631  
Db 1561 GGCACGTATGATGGGAGATTTATTTCCACCTGGGGATGCCATGCAGAGGTTTGGAAAC 1620  
Qy 1632 AAAGAGCATATAAGTGGTATGAGCTTGGGCAACAAGAGAGGACACTTTGCATCTGTCTGG 1691  
Db 1621 AAAGAGCATATAAGTGGTATGAGCTTGGGCAACAAGAGAGGACACTTTGCATCTGTCTGG 1680  
Qy 1692 CAACGCTCACTCTTACAAATGGAATGGAATGGAAGCAAGCTTGGTGGACCCCAAGAA 1751  
Db 1681 CAACGCTCACTCTTACAAATGGAATGGAATGGAAGCAAGCTTGGTGGACCCCAAGAA 1740  
Qy 1752 ACCGGCTTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTTGAAGTTAATCCGAGATGAA 1811  
Db 1741 ACCGGCTTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTTGAAGTTAATCCGAGATGAA 1800  
Qy 1812 GGCCTTGCAGTGTGATGAAGCAAGGTCCTTCTGCTGAGGATGAAGAACTGAGG 1871  
Db 1801 GGCCTTGCAGTGTGATGAAGCAAGGTCCTTCTGCTGAGGATGAAGAACTGAGG 1860  
Qy 1872 GAAAAAGGGGACTGGAGCCAGTTTCAAGCTGTGGCAGCAAGGAAGAGAAATGAAATGCC 1931  
Db 1861 GAAAAAGGGGACTGGAGCCAGTTTCAAGCTGTGGCAGCAAGGAAGAGAAATGAAATGCC 1920  
Qy 1932 TGCAAAAGGAGCTCTTAAACCTGTACCTTACTAGAAAGTTCCCCGAGACAAACAGGATGC 1991  
Db 1921 TGCAAAAGGAGCTCTTAAACCTGTACCTTACTAGAAAGTTCCCCGAGACAAACAGGATGC 1980  
Qy 1992 AGAAGAGGACAGATCAAAATTTCCATCATGCACCCCGGACTCAGCTGTGGCGCGCACACA 2051  
Db 1981 AGAAGAGGACAGATCAAAATTTCCATCATGCACCCCGGACTCAGCTGTGGCGCGCACACA 2040  
Qy 2052 GGGCCCAAAACTGCAGGCTCCGAAATGCACCTGGGCTTGGTGAATCCCAAGGAGGCTGC 2111  
Db 2041 GGGCCCAAAACTGCAGGCTCCGAAATGCACCTGGGCTTGGTGAATCCCAAGGAGGCTGC 2100  
Qy 2112 AAGATTTCGATGTGCCAACGAGACCA 2136  
Db 2101 AAGATTTCGATGTGCCAACGAGACCA 2125

```

DEFINITION Pan troglodytes ASPH gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY405268
VERSION AY405268.1 GI:39761242
SOURCE GSS.
ORGANISM Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 2023)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Fieriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 2023)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Fieriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..2023
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>2023
/gene="ASPH"
/locus_tag="HCM2167"

ORIGIN
Query Match 64.4%; Score 1497.6; DB 9; Length 2023;
Best Local Similarity 74.5%; Pred. No. 0;
Matches 1506; Conservative 0; Mismatches 516; Indels 0; Gaps 0;

QY 115 AGCAAGCATGAGACACACAGATGGGGAAGCGGACTCTCGGACCTTCATCT 174
DB 1 AGACAAAGCATGAGACACACAGATGGGGAAGCGGACTCTCAGGACCTTCATCT 60

QY 175 TCAGTGGTTTATGGTGATTCGATTGCTGGGCGTCTGGACATCTGTAGCTGCTGGT 234
DB 61 TCAGTGGTTTATGGTGATTCGATTGCTGGGCGTCTGGACGCTGTAGCTGCTGGT 120

QY 235 TTGATCTTGTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGCTGATGGTG 294
DB 121 TTGATCTTGTGACTATGAGGAAGTTCTAGNNNNNNNNNNNNNNNNNNNNNNNN 180

QY 295 ATGGAGATTTTGATGGGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACT 354
DB 181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 240

QY 355 CAGAGCAGCAGTCCCGCAGAGAGGCTGAGCACACATGAGCCCGGAGGAGGTC 414
DB 241 CAGAGCAGCAGTCCCGCAGAGAGGCTGAGCACACATGAGCCCGGAGGAGGTC 300

QY 415 CTGPGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 474
DB 301 CTGPGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360

QY 475 TCCATGAAATGGTACACGAGAACATGTTGAGGAGAGAGAGAGAGAGAGAGAG 534
DB 361 TCCATGAAATGGTACACGAGAACATGTTGAGGAGAGAGAGAGAGAGAGAGAG 420

QY 535 CCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 594
DB 421 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 480

```

```

QY 595 GATTGTAGACCCCTGGAACTTGAAGTATCTCATGAAGAAAACGAGCATAGTTACCAGCTGG 654
DB 481 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 540

QY 655 AAGAGACAGTTTCAAGAAGCTCTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGAGAA 714
DB 541 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 600

QY 715 ATCCAGAGTCCAGTGAACCAAGTAGTAGAAGATGAAAGATTGCACCATGATACAGATGATG 774
DB 601 ATCCAGAGTCCAGTGAACCAAGTAGTAGAAGATGAAAGATTGCACCATGATACAGNNNNNN 660

QY 775 TAACATACCAAGTCTATGAGAACCAAGCAGTATATGAACCTCTAGAAAATGAAGGATAG 834
DB 661 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 720

QY 835 AAATCAGAGAGTAACTGCTCCCTCCTGAGGATTAATCTGTAGAAGATTTCACAGTTAATTG 894
DB 721 AAATCAGAGAGTAACTGCTCCCTCCTGAGGATTAATCTGTAGAAGATTTCACAGTTAATTG 780

QY 895 TAGAAGAGTAAAGCATTTTCTGTGGAAGAACAGCAGGAGTAGTACCAACCAAGAAACAAATA 954
DB 781 TAGAAGAGTAAAGCATTTTCTGTGGAAGAACCAAGCAGGAGTAGTACCAACCAAGAAACAAATA 840

QY 955 GAAAAACAGATGATCCAGAACAAAAAGCAAAAGTTTAAAGAAAAGAGCCCTTAAACTTTTAA 1014
DB 841 GAAAAATAGATGATCCAGAACAAAAAGCAAAAGTTTAAAGAAAAGAGCCCTTAAACTTTTAA 900

QY 1015 ATAAATTTGATAGACTATTAAGCTGAACCTGATGCTGAGAAAACCTCCGTAAAGGGT 1074
DB 901 ATAAATTTGATAGACTATTAAGCTGAACCTGATGCTGAGAAAACCTCCGTAAAGGGT 960

QY 1075 GAAAAATTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGCAATACCCCTCAGAGTC 1134
DB 961 GAAAAATTGAGGAAGCAGTGAATGCGTTTAAAGAACTAGTACGCAATACCCCTCAGAGTC 1020

QY 1135 CACGAGCAGATATGGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGAGTA 1194
DB 1021 CACGAGCAGATATGGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGTA 1080

QY 1195 ATGAGGTGCTACCTGGAGCCATCGAGACTACCAAGAGGTGGCCAGCTACCTGATGCTCC 1254
DB 1081 ATGAGGTGCTACCTGGAGCCATCGAGACTACCAAGAGGTGGCCAGCTACCTGATGCTCC 1140

QY 1255 CTGAGACCTGCTGAAGCTGAGTTTGAAGCGTCCGTCAGCAGGCAACATTTCTAGGTC 1314
DB 1141 CTGAGACCTGCTGAAGCTGAGTTTGAAGCGTCCGTCAGCAGGCAACATTTCTAGGTC 1200

QY 1315 ATATGAGAGGTTCCCTGCTTACCTGCGAGAGATTAGTTCAACTATTTTCCCAATGATCTT 1374
DB 1201 ATATGAGAGGTTCCCTGCTTACCTGCGAGAGATTAGTTCAACTATTTTCCNNNNNNNNNN 1260

QY 1375 CTTTAAAAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATAATGACAAATGCAAGA 1434
DB 1261 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1320

QY 1435 AAGTTTATGAAGAGGTGCTGAGTGCACCTTAATGATGGCTTGTCTAAGTCCATTATG 1494
DB 1321 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1380

QY 1495 GCTTCATCTGAAAGGACAGAAACAAAATTCGTGAGAGATCCCONATTTTAAAGAGAGGAA 1554
DB 1381 GCTTCATCTGAAAGGACAGAAACAAAATTCGTGAGAGATCCCONATTTTAAAGAGAGGAA 1440

QY 1555 TAGAATCCGAGAGATCCCTGCACTGATGAGGAGATTTTATTTCCACTCTGGGGATGCCA 1614
DB 1441 TAGAATCCGAGAGATCCCTGCACTGATGAGGAGATTTTATTTCCACTCTGGGGATGCCA 1500

QY 1615 TGCAGAGGTTGGGAACAAAGAGGCATATAAGTGGTATGAGCTTGGGSCACAAGAGAGGAC 1674
DB 1501 TGCAGAGGTTGGGAACAAAGAGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1560

```

```

Qy 1675 ACTTGTGATCTGTCTGCGAACGCTCACTCTACAATGTAATGGACTGAAAGCAGACGCTT 1734
Db 1561 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1620
Qy 1735 GGTGAGCCCAAGAAACGGGCTACACAGATTAGTAAGTCTTTAGAAAGAACTGGA 1794
Db 1621 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1680
Qy 1795 AGTTAAATCCGAGATGAAGGCTTCAGTGTATGGAATAAGCAAGAGTCTCTTCTGCTG 1854
Db 1681 AGTTAAATCCGAGATGAAGGCTTCAGTGTATGGAATAAGCAAGAGTCTCTTCTGCTG 1740
Qy 1855 AGGATGAAACCTCAGGAGAAAGGGGACTGGACCCAGTTCACGCTGTGGCAGCAAGAA 1914
Db 1741 AGGATGAAACCTCAGGAGAAAGGGGACTGGACCCAGTTCACGCTGTGGCAGCAAGAA 1800
Qy 1915 GAAGAATGAATAATGCTGCAAGAGCTCTTAAACCTGTACCTTACTAGAAAGTTCC 1974
Db 1801 GAAGAATGAATAATGCTGCAAGAGCTCTTAAACCTGTACCTTACTAGAAAGTTCC 1860
Qy 1975 CCGAGACAAACAGGATGAGAGAGGACAGATCAATATTCATCATGCACCCCGGACTC 2034
Db 1861 CCGAGACAAACAGGATGAGAGAGGACAGATCAATATTCATCATGCACCCCGGACTC 1920
Qy 2035 ACGTGTGCGCCGACACAGGCGCCCAAACTCGAGCTCCGAATGCACCTGGGCTTGTGA 2094
Db 1921 ACGTGTGCGCGACACAGGCGCCCAAACTCGAGCTCCGAATGCACCTGGGCTTGTGA 1980
Qy 2095 TTCCCAAGGAGGCTGCAAGTTGATGTGTCACAGGACCA 2136
Db 1981 TTCCCAAGGAGGCTGCAAGTTGATGTGTCACAGGACCA 2022

RESULT 3
AK030293
LOCUS AK030293 4406 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus 11 days pregnant adult female ovary and uterus cDNA,
RIKEN full-length enriched library, clone:5031409C19
product:aspartate-beta-hydroxylase, full insert sequence.
ACCESSION AK030293
VERSION AK030293.1 GI:26326276
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci,P. and Hayashizaki,Y.
2
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
REFERENCE
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
REFERENCE
AUTHORS Carninci,P., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuana,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanabe,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Tanaka,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipillar sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
PUBMED

```

```

REFERENCE
AUTHORS FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/
FEATURES
source
Location/Qualifiers
1..4406
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:5031409C19"
/db_xref="taxon.10090"
/clone="5031409C19"
/sex="female"
/tissue_type="ovary and uterus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="11 days pregnant adult"
177..2402
/notes="unnamed protein product; aspartate-beta-hydroxylase
(MGD|MG1:1914186, GI|NW_023066, evidence: BLASTN, 99%,
match=4404)
putative"
/codon_start=1
/db_xref="GI:26326277"
/translation="MAPRKNKAGCGGNSSSGSGSGSGSPSTGSSGSSSPGABRE
AKHGDFVDDKAVLLGLKERSPTFPPEEAETHAELEQAPGADIQNVDEPKQ
IQSLLOESVHTDHLADGLAGPQPEVEDFLTVTDSDDFDELPQTVHEELEDYTH
VEPTASQNPNDMEEMTNEQNSDPSAEVTDAGVLLPHAEVRHQDYDEVPYPSSEH
GVAISNTIDSSIISEINVASVERQDTPPVKKKKPLKNFKDFTIKAELEAEKL
RRKGEANTVAEELVRKYQPQPRARYGAQCEDLAEQRQNEVLRAIEYQEAEL
LDLPAPDVLVNLKRSEKQITLQRLVQLFPPSDTLKNDLGVGLL
GNDNSAKVVEEVLNTPNDPAKHYGFILKAQNKISEIPYLKGEAKQAPWMTTGG
DYFPHLGDAMQVRGNKEAYKVELGHRKFHAFSWQSLNVNGLKAPQWMTTGGY
TELVKSLERNWLIRDEGLMVKAGLFLPEDENLRKGDWSQTLKQCRKNENAC
KGAPKCALLKEFSETTGCRRGOIKYSIMHPGTHVMPHTGTGTCRLRMHGLVLPKRG
CKTRCANETTWEEGKVLIFDDSFEEHVWQDASSFRLLIFIVDVWHPPLITPQRRSLPA

```

```

4 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 4406)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/
FEATURES
Location/Qualifiers
1..4406
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:5031409C19"
/db_xref="taxon.10090"
/clone="5031409C19"
/sex="female"
/tissue_type="ovary and uterus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="11 days pregnant adult"
177..2402
/notes="unnamed protein product; aspartate-beta-hydroxylase
(MGD|MG1:1914186, GI|NW_023066, evidence: BLASTN, 99%,
match=4404)
putative"
/codon_start=1
/db_xref="GI:26326277"
/translation="MAPRKNKAGCGGNSSSGSGSGSGSPSTGSSGSSSPGABRE
AKHGDFVDDKAVLLGLKERSPTFPPEEAETHAELEQAPGADIQNVDEPKQ
IQSLLOESVHTDHLADGLAGPQPEVEDFLTVTDSDDFDELPQTVHEELEDYTH
VEPTASQNPNDMEEMTNEQNSDPSAEVTDAGVLLPHAEVRHQDYDEVPYPSSEH
GVAISNTIDSSIISEINVASVERQDTPPVKKKKPLKNFKDFTIKAELEAEKL
RRKGEANTVAEELVRKYQPQPRARYGAQCEDLAEQRQNEVLRAIEYQEAEL
LDLPAPDVLVNLKRSEKQITLQRLVQLFPPSDTLKNDLGVGLL
GNDNSAKVVEEVLNTPNDPAKHYGFILKAQNKISEIPYLKGEAKQAPWMTTGG
DYFPHLGDAMQVRGNKEAYKVELGHRKFHAFSWQSLNVNGLKAPQWMTTGGY
TELVKSLERNWLIRDEGLMVKAGLFLPEDENLRKGDWSQTLKQCRKNENAC
KGAPKCALLKEFSETTGCRRGOIKYSIMHPGTHVMPHTGTGTCRLRMHGLVLPKRG
CKTRCANETTWEEGKVLIFDDSFEEHVWQDASSFRLLIFIVDVWHPPLITPQRRSLPA

```

polyA_signal		I"	4387..4392		DB 3; Length 4406;				
		/note="putative"		79.4%; Pred. No. 0;					
polyA_site		4406		Matches 1834; Conservative 0; Mismatches 380; Indels 96; Gaps 6;					
		/note="putative"							
ORIGIN									
Query Match 62.4%; Score 1450; DB 3; Length 4406;									
Best Local Similarity 79.4%; Pred. No. 0;									
Matches 1834; Conservative 0; Mismatches 380; Indels 96; Gaps 6;									
QY	15	GCCAGCGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCGGCTCCGGCAGCGGT	74						
DB	195	GCCAAAGCGCGCGCGCAACAGCAGCAGCGGCGCAGCGGCTCCGGCAGCGGT	254						
QY	75	AGCAGGATCGGGCGAGCAGC-----AGCCCGGGCGCGGAGAGGACAAAG	122						
DB	255	AGCCCGAGCAGCGGTAGCAGCGGTAGCAGCAGGTAGCCCGGGGCTCGGAGAGGCAAG	314						
QY	123	CATGGAGGACACAAGATGGAGGAAAGCCGACTCTCGGAACTTCATTTCACTGG	182						
DB	315	CACGGAGGACACAAGAAATGGAGGAAAGGAGGATTCAGAGGGTCTTTTTCACATGG	374						
QY	183	TTTATGGTGAATTGCATTGCTGGGCGCTCGACATCTGTAGCTGTGTTGGTTGATCTT	242						
DB	375	TTCATGTGTCATTGCTGCGGCTCTGGACATCTGTGGCTGTGTTGGTTGACTTG	434						
QY	243	GTTCACATCAGGAGAGTCTTAGGAAACTAGGAATCTATGATGCTGATGTGATGGAGAT	302						
DB	435	GTGATTTAGCAAGAGTCTTAGGAAACTAGGAGTCTATGATGCGATGCGATGGAGAC	494						
QY	303	TTTGATGTGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGGCCA	362						
DB	495	TTTGATGTGATGATGCCAAAGTTTATTAGGCTTAAAGAAAGATCTCTTCTGAGAGG	554						
QY	363	GCAGTCCCGCC---AGAAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGAGTCTCTGTG	419						
DB	555	ACATTTCCACCGGAGGAGGAGGAGGAGACTCACGTGAGCTGGAGGAGGAGGCGCCCTGAG	614						
QY	420	GAGCAGAAACCCAGAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCAT	479						
DB	615	GGGCGACACATCCAGAAATGTTGAAGATGAAGTAAAGGAAACAAATTCAGTCCCTTCTCAG	674						
QY	480	GAAATGGTATACCCAGACATGTTGAGGAGAGAGACTTGCAACAGAGAGATGAGCCACA	539						
DB	675	GAATCAGTACACACAGACCATTG-----ACCTTGAAGCAGATGGGCTGGCA	719						
QY	540	GGAGAACCAACAGAGGATGATGAGTTTCTTATGGGAGTGTAGATGATGATGATTT	599						
DB	720	GGAGAACCAACAGCGGAGGTTGAGGATCTCTCTCAGTGACCGACAGTATGACAGATTT	779						
QY	600	GAGACCTTGGAACTTGAAGTATCTCATGAAGAAACCGAGCAGTATGTTACCGTGAAGAG	659						
DB	780	GAGGACCTTGAACCCGGGACAGTTTCATGAAGAAATCGAGGATCTTACCATGTGAAGAC	839						
QY	660	ACAGTTTCAAGAGCTGTAATCAGGATATGGAAGAGATGATGTCTGAGCAGGAAATCCA	719						
DB	840	ACAGCATCGCAGAACCATCCAAATGACATGGAAGAGATGAGAAATGAGCAGGAAATTC	899						
QY	720	GATTCAGTGAACAGTAGTAGAAGATGAAGATGTCACCATGATACAGATGATGAACA	779						
DB	900	GATCCAGTGAAGCAGTGAACATGACAGTGTGCTGTGCTGCCCATGACAGGAGATGA	959						
QY	780	TACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTTGAAGAAATGAGGAGTAGAATC	839						
DB	960	CACCAAGACTATGATGAAC---CAGTATATGAACCTCAGAGCATGAAGGGTTCGGGATT	1016						
QY	840	ACAGAAGTAACTGCTCCCTCAGGATATATCTGTGTAAGATTCACAGGTAATGTGAAG	899						
DB	1017	TCA-----GATTAACCATAGATGATTCAGCATATCTCAGCAATATCTCAGAA	1055						
QY	900	GAAATAGCATTTTCTGTGTGAAGAACAGCAGGAGAGTACCAACACAGAAACAAATAGAAA	959						

DB	1056	GAATTAATGTGCGCTCTCTGTGAAGAACAGCAGACACACACC-----	1099		
QY	960	ACAGATATCCAGAACAAAGAGTTTAAAGAAAGAGCCTAAACTTTTAAATAAA	1019		
DB	1100	-----AGTTAAGAAAGAGAGCCTAAACTTTTGAACAA	1133		
QY	1020	TTTGATAAGACTATTAAAGCTGAACCTTGATGCTGCGAGAAACCTCCGTAAAGGGGAAA	1079		
DB	1134	TTTGATAAGCAATTAAGCTGAGCTGGATGCTGCGAGAAAGCTCCGGAAGGGGTAAA	1193		
QY	1080	ATTGAGAGACAGCTGAATCATTTAAAGAACTTAGTAGCAAAATACCTCAGAGTCCACCA	1139		
DB	1194	ATTGAGAGACAGCTGAACCATTTGAAGAACTGGTTGMAAGTACCCTCAGAGCCACCA	1253		
QY	1140	GCAAGATATCGGAAGCGCGAGCTGAGATGATTTGGCTGAGAGAGAGAGTATGAG	1199		
DB	1254	GCAAGATATCGCAAGCGCGAGTGTGAAGATGACTTTGGCAGAGAGACAGAGCAATGAG	1313		
QY	1200	GTGCTACGTGGAGCCATCGAGACTTACCAAGAGGTGGCGAGCTTACCTGATGCTCCGCA	1259		
DB	1314	GTCTGCGCAGGGCCATCGAGACTTACCAGGAGGCGAGCCGCTGCTGATGCCCTACA	1373		
QY	1260	GACCTGCTGAAGCTGAGTTTGAAGGCTGCTCAGACAGGCAACAAATTTCTAGGTCAATG	1319		
DB	1374	GACCTGCTGAAGCTGAGCTTGAAGCGAAAGGTGCGAAACGGCAGCAGTTTCTAGGTCAATG	1433		
QY	1320	AGAGGTTCCCTGCTTACCTCGCAGAGATTTAGTTCAACTTTTCCCAATGATCTTCCCTTA	1379		
DB	1434	AGAGGTTCTCTACTTACCTTCCAGAGACTAGTTTCACTGTTTCCCTAGTATCTACTTA	1493		
QY	1380	AAAAATGACCTTGGCTGGGATACCTCTTGATAGAGATATGACATTCGAAAGAAAGTT	1439		
DB	1494	AAAAAACACCTTGGCTGGGATACCTCTTGTGGGAGACCAACAGCTGCGCAAGAGGTT	1553		
QY	1440	TATGAGAGGTGCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTATGGCTTC	1499		
DB	1554	TACGAGAGGTGCTAAATGTGACACAAATGATGGCTTCGCTAAAGTGCTATTCGCGCTTC	1613		
QY	1500	ATCTGTAAGGCAACAGAAATGCTGAGAGCATCCCATATTTTAAAGGAAGGAATAGAA	1559		
DB	1614	ATCTGTAAGGCAACAGAAATATCTGAGAGCATTCCTACTTTAAAGGAAGGAATCGAA	1673		
QY	1560	TCCGAGATCCTGGCAGTATGATGGGAGATTTTATTTCCACCTGGGGATGCCATGAG	1619		
DB	1674	TCTGGGACCTGGCAGGATGATGGCGGTTTTACTTCCACTTGGGGATGCCATGAG	1733		
QY	1620	AGGTTGGGAAACAGAGGCATATAAGTGGTATGAGCTTGGGCAACAGAGAGACATTT	1679		
DB	1734	AGGTTGGGAAACAGAGGCATATAAGTGGTATGAACTTGGGCAACAGAGAGACATTT	1739		
QY	1680	GCATCTCTGTGGCAACGCTCACTCTCAATGTGAATGGACTGAAAGCAGACGCTTGGTGG	1739		
DB	1794	GCCTCTCTGTGGCAGCGTTCCCTCTCAATGTGAATGGTCTGAAAGGCTCAGCCGCTGG	1853		
QY	1740	ACCCCAAAAGAAAGCGGCTACACAGAGTTTAGTAAAGTCTTTAGAAAGAACTCGAAGTTA	1799		
DB	1854	ACACCCAGGAGACTGGCTTACACAGAGCTAGTGAAGTCTTTAGAGAGAACTTGAAGTTA	1913		
QY	1800	ATCCGAGATGAAGCGCTTCGATGATGATTAAGCCAAAGTCTCTTCTGCTGCTGAGAT	1859		
DB	1914	ATCCGATGAAGCGCTCATGTGATGATTAAGCCAAAGGCTCTCTTCTGCTGCTGAGAC	1973		
QY	1860	GAAGAACTGAGGAAAAAGGGGACTGGAGCCAGTTTCACTGTGGCAGCAAGGAAGA	1919		
DB	1974	GAAGAACTTCCGAGAAAGGGGACTGGAGCCAGTTTCACTGTGGCAGCAAGGAAGA	2033		
QY	1920	AATCAAAATCCCTGCAAGAGCTCTTAAACCTGTACTTACTAGAAAAGTTTCCCGAG	1979		
DB	2034	AATGAGAACTGCTGTAAGAGCGCTTAAGACCTGTCTTACTAGAAAAGTTTCCGAA	2093		
QY	1980	ACACAGATGACAGAGGAGCAGATCAATTTCCATCATGACCCCGGACTCAGCTG	2039		
DB	2094	ACAACAGATGACAGAGGAGCAGATCAATTTCCATCATGACCCCTGGAACTCATGTG	2153		







RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama  
Kanagawa 230-0045, Japan [E-mail: genome-resgsc.riken.jp,  
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216]  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.jp/  
URL: http://genome.gsc.riken.jp/  
Location/Qualifiers  
1. .2854  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:9530097J16"  
/db\_xref="taxon:10090"  
/clone="9530097J16"  
/sex="male"  
/tissue type="urinary bladder"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
263. .2239  
/note="unnamed protein product; aspartate-beta-hydroxylase  
(MGD|GI:1914186, GB|NM\_023066, evidence: BLASTN, 99%,  
match=4404)  
putative"  
/codon\_start=1  
/protein\_id="BAC29171.1"  
/db\_xref="GI:26330882"  
/translation="MVIALGLGVMTSVAVVMFVLDVLYEVLVGLGVVDADGDGDFDVID  
AKVLGLKERSPTFPPEAEAEHAELEQAGADQNVDEVEKEIQSLQDSVS  
HTDHLADGLAGEPQSEVDPLTVTSDSDREDELPGETVHEIEDTVHEDTASQNH  
PNDMEETMEQENSEVEHQDYDLYVPYSPSEHEGVAISDNTIDDSIISEENVASVE  
EQQDTPPVKKKPKULNPKDKTIAELBAEKLKRGKLEAVNAFEELVRKYPQSPRR  
ARYKQAQCEDLAEQKQNSVLLRAETIYQEAADLPDADTVLKSLSKRSEKQFLG  
HMRGLSLDLQRLVQFSPEDTLNGLDGYGLLLGNDSDAKVYVEEVLNTPNDQFAK  
HYRTLKAAQKISGIVLYKEGIESGDGPTDGRFVFLGLDAMQVGNKEAYKWLGLG  
HGRGHFAVNQSLPVNGLKAQPMWTRETGYTELKSLERNKWLIRDBGLVMVMDKA  
KGLFLPEDENLRKGDWQFTLWQQRKNENACKGPKTCALLEKSETTGTGCRGOIK  
YSIMHPGTHVMPHTGFTNCRLMHGLVLPKGGKLRICANETRTWEEGKVLIFDDSFPE  
HEVMQDASSFLRIFIVDVHWPBLTPQRRSLPAI"

QY 471 CTTCTCCATGAATGGTACACGACGACATGTTGAGGAGAGAGACTTGCACACAGAGAT 530  
Db 551 CTTCTTCAGGATCAAGTACACAGACCATG-----ACCTTGAGCAGAT 595  
QY 531 GGAACCCACAGAGAACACACACAGAGAGATGATGATTTCTTATGGGACCTGATGATGAT 590  
Db 596 GGGCTGGCAGAGAACACACAGCGAGGTTGAGGACTTCTCCACAGTGCACGACAGTAT 655  
QY 591 GATGATTTGAGACCCCTGGAACTCTGAAGTATCTCATGAGAAACCCGAGCATAGTTACAC 650  
Db 656 GACAGATTTGAGGACCTGGAACTCCCGGACAGTTTCATGAGAAATTCAGAGATACCTACCAT 715  
QY 651 GTGGAAGAGACAGTTTCACAAAGACTGTATCAGGATATGGAAGAGATGATGTTGAGCAG 710  
Db 716 GTGGAAGACACACAGCATCGCAGAACCATCCAAATGACATGGAAGAGATGACGATGAGCAG 775  
QY 711 GAAATCCAGATTCAGTGAAACCATGATAGTATGAGAGATGAAAGATTGACCATGATACAGAT 770  
Db 776 GAAATTCAGAG-----787  
QY 771 GATGTATACATACCAAGTCTATGAGGAAACAGCAGTATATGAACTCTAGAAAATGAGGG 830  
Db 788 GAAATGATGACCAAGCATATGATGAAC---CAGTATATGAAACCTCAGAGCATGAAAGG 844  
QY 831 ATAGAAATTCAGAAAGTAACTGCTCCCTGAGGATATCTCTGAGAAAGATTACAGGTA 890  
Db 845 GTCGCGATTTCA-----GATAACCATATGATATCCAGCATA 883  
QY 891 ATTCTAGAAGATGAGCATTTTCTCTGGAAGAACAGCAGGATGATACCCAGAGAAACA 950  
Db 884 ATCTCAGAGAAATTAATGTGCTCTCTGTTGAGAAACAGACAGACACACAC-----936  
QY 951 AATGAAAAACAGATATCCAGAACAAAAGGTTAAGAAAAGAGGCTTAACTT 1010  
Db 937 -----AGTTAAGAAAGAGGCTTAACTT 961  
QY 1011 TTAATTAATTTGATTAAGACTATTAAGCTGAACCTGATGCTCGAGAAAACCTCGTAAA 1070  
Db 962 CTGAACAAATTTGATTAAGACAATTAAGCTGAGCTGGATGCTCGAGAAAAGTTCGGAAA 1021  
QY 1071 AGGGGAAAAATTCAGGAAGCAGTGAATGCTTTAAAGAACTAGTAGCAAAATACCTCAG 1130  
Db 1022 AGGGTAAAAATTCAGGAAGCAGTGAACGATTTGAAGAACTGTTGMAAGTACCTCAG 1081  
QY 1131 AGTCCACGACAGATATGGGAAGGCGCAGTGTGAGATGATTTGCTGAGAGAGAGA 1190  
Db 1082 AGCCACAGACAGATATGGCAAGCGCAGTGTGAGATGATTTGCTGAGAGAGAGAGAGA 1141  
QY 1191 AGTATGAGTGCTAGCTGGAGCCATCGAGACCTTACCAAGAGTGTCGAGCCTACCTGAT 1250  
Db 1142 AGCAATGAGTTCTGCGCAGGCGCCATCGAGACCTTACAGAGGAGCGCCACCTGCTGAT 1201  
QY 1251 GTCCCTGCAGACTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAAATTTCTA 1310  
Db 1202 GCCCTACAGACCTGCTGAAGCTGAGCTTGAAGCGAAGTTCGAAACGCGCAGCTTTCTA 1261  
QY 1311 GGTCAATGAGAGTTTCCCTGCTTACCTGCGAGAGATTAAGTCAATTTTCCCAATGAT 1370  
Db 1262 GGTCAATGAGAGTTTCTCTACTTACCTCCAGAGACTAGTTCACTGTTCTCCTAGTAT 1321  
QY 1371 ACTTCTTAATAAATGACCTTGGCTGGGATACCTCTTGTATGAGAGATATGACAAATGCA 1430  
Db 1322 ACTACTTTAAAAACGACCTTGGCTGAGTATACCTCTTGTGGAACACACAGAGTCC 1381  
QY 1431 AAGAAAGTTTATGAGAGGTGCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCAT 1490  
Db 1382 AAGAAAGTTTATGAGAGGTGCTAATATGAGACCAATGATGGCTTCGCTAAAGTGCAT 1441  
QY 1491 TATGGCTTCACTCTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCATATTTAAAGAA 1550  
Db 1442 TAGGGCTTCACTCTGAAGGCACAGAAACAGATATCTGAGAGCATTCCTTACTTTAAGAA 1501

QY 1551 GGAATAGANTCCGAGATCTGCACTGATGATGGAGATTTTATTTTCCACCTGGGGAT 1610  
Db 1502 GGAATCGANTCTGGGAGCCTTGCAAGATGATGCGGGTTTACTTTCCACTTGGGGAT 1561  
QY 1611 GCCATGACAGAGGTGGGAAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAACAAGAGA 1670  
Db 1562 GCCATGACAGAGGTGGGAAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAACAAGAGA 1621  
QY 1671 GGAACATTTTCATCTGTCTGGCAAGCTCACTCTACATGTGAATGAGTGAAGACACACAG 1730  
Db 1622 GGAATTTTCCTCTGTCTGGCAGCGTTCCCTCTACAATGTGAATGCTTGAAGGCTCAG 1681  
QY 1731 CTTTGTGGACCCCAAAAGAAACGCGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAAC 1790  
Db 1682 CCGTGTGGACACCCAGGAGAGCTGGCTACACAGAGCTAGTAAAGTCTTTAGAGAGAAAC 1741  
QY 1791 TGGAGTTAATCCGAGATGAAGGCTTTCAGTATGATGAATGAAGCAAAAGGCTCTTCTCTG 1850  
Db 1742 TGGAGTTAATCCGATGAAGGCTTTCAGTATGATGAATGAAGCAAAAGGCTCTTCTCTG 1801  
QY 1851 CCTGAGATCAAAACCTGAGGAAAGAGGAGCTGGAGCCAGTTTCAAGCTGTGGCAGCAA 1910  
Db 1802 CCTGAGACGAAACCTTCCGGAGAGAGGCGAGCTGGAGCCAGTTTCAAGCTGTGGCAGCAA 1861  
QY 1911 GGAAGAAAGAAATGAAATGCTTCAAGAGAGCTCTTAAACCTGTACTTACTTAGAAAG 1970  
Db 1862 GGAAGAAAGAAATGAGATGCTTAAAGAGGCGCTAAGACCTGTGCTTTACTAGAAAG 1921  
QY 1971 TTCCCGGAGACACAGATGACAGAGAGACAGATCAATATTCATCATGCAACCCCGG 2030  
Db 1922 TTTTCCGAAACACAGATGACAGAGAGACAGATCAATATTCATCATGCAACCCCGG 1981  
QY 2031 ACTCAGCTGTGGCGCACACAGGCGCCACAACTGCGAGGCTCCGAATGCACTTGGGCTT 2090  
Db 1982 ACTCAGTGTGGCGCATACAGGACCCACAACTGCGAGGCTCCGAATGCACTTGGGCTT 2041  
QY 2091 GTGATTTCCAAAGAGGCTGCAAGATTCGATGTGCCAACGAGACCCAGGAGGAA 2150  
Db 2042 GTGATTTCCAAAGAGGCTGCAAGATTCGATGTGCCAACGAGACCCAGGAGGAA 2101  
QY 2151 GGAAGTGTCTCATCTTTGATGACTCTTTGAGCAGCAGGTATGGAGGATGCTCATCT 2210  
Db 2102 GGAAGTGTCTCATCTTTGATGACTCTTTGAGCAGCAGGTATGGAGGATGCTCATCT 2161  
QY 2211 TTCCGCTGATATTCATGCTGATGTGTGGATCCGAACTGACACACAGCAGAGAGC 2270  
Db 2162 TTCCGCTGATATTCATGCTGATGTGTGGATCCGAACTGACACACAGCAGAGAGC 2221  
QY 2271 AGCTTCCAGCAATTTAGCATGAATTCATCAAGCTTGG 2309  
Db 2222 AGCTTCCGCAATTTGAAGGCACTGACGAGACTTGG 2260

## RESULT 6

AY405269 Mus musculus ASPH gene, partial sequence, linear GSS 16-DEC-2003  
LOCUS AY405269 genomic survey sequence.  
DEFINITION AY405269, (bases 1 to 2153)  
ACCESSION AY405269, GI:39761243  
VERSION AY405269.1  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 2153)  
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,  
Todd, M.A., Tanenbaum, D.M., Clivello, D.R., Lu, P., Murphy, B.,  
Perrera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
Adams, M.D. and Cargill, M.  
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)

```
PUBMED 14671302
REFERENCE 2 (bases 1 to 2153)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Ranenbaum,D.M., Civello,D.R., Lu,P., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
    source      Location/Qualifiers
        1..2153
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            <1..>2153
            /gene="ASPH"
            /locus_tag="HCM2167"
    gene
    ORIGIN
        Query Match      54.4%; Score 1264; DB 9; Length 2153;
        Best Local Similarity 69.5%; Pred. No. 5.6e-313;
        Matches 1465; Conservative 0; Mismatches 638; Indels 6; Gaps 2;

QY 32 CAAGAGCAGCGGCAACAGCAGCAGCGGCTCCGGCAGCGGTAGCACGAGTGGCGGCGAG 91
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92 CAGCAGCCCCGGGCCCCGGAGAGAGACAAGCATGGAGGACACAAGAAATGGAGGAAGG 151
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 108 CAGTAGCCCCGGGCTCGAGAGAGGCAAGACAGGAGGACACAAGAAATGGAGGAAGG 167
QY 152 CGGACTCTCGGAACCTCATCTTCACGTGGTGTATGTTGATTCGATTCGTCGGCGCTG 211
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 168 AGGGAATTCAGGAGGGTCTTTTTCACATGGTTCATGGTCAATTCGATTCGAGGAGT 227
QY 212 GACATCTAGTGTGCTGTTGTTGATCTTGTGTTGACTATCAGGAAGTTCCTAGGAAACT 271
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 228 GACATCTGTGCTGTGTTGTTGATCTTGTGTTGACTTCGATTCAGGAAGTTCCTAGN 287
QY 272 AGGAATCTATGATGCTGATGATGAGATTTTGAATGATGATGATGATGATGATGATGAT 331
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 288 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 347
QY 332 AGGACTTAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCC---AGAGAGGCTGAGCC 388
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 348 NNGCCTTAAGAAAGATCTCTTCTGAGAGACATTTCCACGGAGGAGGAGGAGAGAC 407
QY 389 ACACACTGAGCCCGAGGAGCAGGTTCTGTGGAGCAGAACCCAGAAATATCGAAGATGA 448
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 408 TCACGCTGAGCTGGAGGAGCAGGCCCTGAGGGGCGACACATCCAGAAATGTTGAAGATGA 467
QY 449 AGCAAAAGAACAAATTCAGTCCCTTCTCATGAAATGTTACACCGCAACATGTTGAGGG 508
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 468 AGTAAAGGAACAAATTCAGTCCCTTCTTCAGGAATTCAGTACACACAGACCATGCTTGT 527
QY 509 AGAAGCTTGCAACAAGAGATGGACCCACAGGAGAACCAACAAGAGGATGATGAGTT 568
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 528 TGAAGAC---CTTGAAGCAGATGGCTGGCAGGAGAACCAACGCGGAGGTTGAGACTT 584
QY 569 TCTTATGGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 628
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 585 CCTCACAGTACCGACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 644
QY 629 AGAAACCGGACATAGTACCAGTGGAGAGACAGTTTTCACAGACTGTAATCAGGATAT 688
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 645 AGAAATCGAGGATACCTTACCATGTGGAAGACACACATCGCAGAACCATCCAAATGACAT 704
QY 689 GGAAGAGATGATGCTGAGCAGGAAATCCAGATTCAGATTCAGATTCAGATTCAGATTCAG 748
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 705 GGAAGAGATGACGAATGACGAGGAAATTCAGNNNNNNNNNNNNNNNNNNNNNNNNNNNN 764
QY 749 AAGATTGCACCATGATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 808
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 765 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 824
QY 809 TGAACCTCTAGAAATGAAGGATAGAAATCACAGAAGTAACTGCTCCCTCCTGAGGATAA 868
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 825 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 884
QY 869 TCCTGTAGAAGATTACAGGTAAATTGTGAAGAAGTAAAGCATTTTCTCTGTGGAAGACA 928
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 885 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 944
QY 929 GCAGGAAGTACACACAGAAACAAATAGAAAAACAGATGATCCAGAACAAAAAGCAAAAGT 988
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 945 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1004
QY 989 TAAGAAAAAGAGCGCTAAACTTTTAAATAAATTTGATAAGACTATTAAGAGCTGAACTTGA 1048
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1005 TAAGAAGAAGAGCGCTAACTTCTGAACAAATTTGATAGACAAATTAAGCTGAGCTGGA 1064
QY 1049 TGCTGCAGAAAACTCCGTAAAAAGGGGAAAAATTTGAGGAAGCAGTGAATGCAATTTAAAGA 1108
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1065 TGCTGCAGAAAAAGCTCCGGAAGGGGTAAAAATTTGAGGAAGCAGTGAAGCAATTTGAAGA 1124
QY 1109 ACTAGTACGCAATACCTTCAGAGTCCAGAGCAAGATATGGGAAGCGCAGTGTGAGGA 1168
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1125 ACTGGTTGCAAGTACCTTCAGAGCCACAGACAGATATGGCAAGCGCAGTGTGAAGA 1184
QY 1169 TGATTTGGCTGAGAAGAGGAGAGTAAATGAGGTGCTACGTGGAGCCATCGAGACCTACCA 1228
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1185 TGACTTGGCAGAGAGAGAGAGCAATGAGGTTCTTGGCAGGGCCATCGAGACCTACCA 1244
QY 1229 AGAGGTGGCAGCCTACCTGATCTCCCTGAGACCTGCTGTAAGCTGAGTTTGAAGCGTGG 1288
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1245 GGAGGAGCGCGACCTGCTGATGCCCTACAGACCTGGTGAAGCTGAGCTTGAAGCGAAG 1304
QY 1289 CTGAGCAGGCAACAATTTCTAGTTCATATGAGAGGTCCCTCTTACCTCGCAGAGATT 1348
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1305 GTCGGAACCGCAGCAGTTCTAGGTACATGAGAGGTTCTTACTTACCTCCAGAGACT 1364
QY 1349 AGTTCAACTATTTCCCAATGATATCTTCTTAAAAAATGACCTTTGGCGTGGGATACCTCTT 1408
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1365 AGTTCAACTGTTCCCTAGTATACCTTAAAGACGACCTTTGGCGTAGGATACCTCTT 1424
QY 1409 GATAGGAGATATGACAAATGCAAGAAAGTTTATGAAAGAGTCTGAGTGTGACACCTAA 1468
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1425 GTTGGGAGACAACAGCAGTCCCAAGAGGTTTACGAAGAGGTGCTAAATGTGACACCAA 1484
QY 1469 TGATGGCTTTGCTAAAGTCCATTATGGCTTCATCTGAAAGGCACAGAAACAAATTTGCTGA 1528
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1485 TGATGGCTTCGCTAAAGTGCATTTACGGCTTCATCTGAAAGGCACAGAAACAAAGATATCTGA 1544
QY 1529 GAGCATCCCATATTTAAAGGAAGGAATAGAAATCCGAGATCCTGGCATCTGATGATGGAG 1588
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1545 GAGCATTTCCCTACTTTAAAGGAGGAATCGAATCTGGGGACCTTGGCAGCGATGATGCCG 1604
QY 1589 ATTTTATTTCCACTGCGGGGATGCGATGAGAGGTTGGGAAACAAAGAGGCATATAAGTG 1648
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1605 GTTTTACTTCCACTTGGGGATGCGATGCGAGGGTCGGGAAACAAAGAGGCATATAAGTG 1664
QY 1649 GTATGAGCTTTGGGCACAGAGAGGACACTTTTGGCATCTGTCTGGCAACGCTCCTCTACAA 1708
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1665 GTATGAACTTTGGGCACAGAGAGGACATTTTGGCTCTGTCTGGCAGGTTCCCTCTACAA 1724
QY 1709 TGTGAATGGACTGAAAGCAGAGCCTTGGTGGACCCCAAAAGAAACGCGGTACACAGAGTT 1768
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1725 TGTGAATGTTCTGAAGGCTCAGCGCTGGTGGACACCCAGGGAGACTGGCTACACAGAGCT 1784
QY 1769 AGTAAAGCTTTAGAAAGAAACTGGAAGTTAATCCGAGATGAAGGCTTGCAGTGTAGGA 1828
Db  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1785 AGTGAAGCTTTTAGAGAGAAACTGGAAGTTAATCCGATGAAGGCTCATGGTGTAGGA 1844
QY 1829 TAAAGCCAAAGGTTCTTCTCTGCTGAGGATGAAAAACCTGAGGGGAAAAAGGGGAGCTGGAG 1888
```

Db 1845 TAAAGCCAGGGTCTTTCTGCTGAGGACGAAACCTTCGGGAGAGGGCGACTGGAG 1904  
QY 1889 CAGTTCACGCTGTGGCAGCAAGAGAAAGAAATGAAATGCTTCGCAAGAGGAGCTCTAA 1948  
Db 1905 CAGTTCACACTGTGGCAGCAAGAGGAGGAATGAGAAATGCTGTGAAAGGAGCGCTAA 1964  
QY 1949 AACCTGTACCTTACTAGAAAAGTTCCCGCAGACACAGAGGATGAGAGAGGAGCAGATCAA 2008  
Db 1965 GACCTGTGCTTTACTAGAAAAGTTTCCGAAACACAGGATGAGAGAGGAGCAGATCAA 2024  
QY 2009 ATATTCCATCATGCCCGGAGCTACAGTGTGGCCGACACAGGCGCCACAACTGCGAG 2068  
Db 2025 ATACTCCATCATGCCCGCTGGAACCTCATGTGTGGCGGATACAGGACCCACAACTGCGAG 2084  
QY 2069 GCTCGGATGCACTGGGCTTGGTATTCCTCCAGAGAGGCTGCAAGATTGCGATGCGCAA 2128  
Db 2085 GCTCGGATGCACTGGGCTTGGTATTCCTCCAGAGAGGCTGCAAGATTGCGATGCGCAA 2144  
QY 2129 CGAGACCAC 2137  
Db 2145 TGAGACCAC 2153

## RESULT 7

BM806632 1050 bp mRNA linear EST 05-MAR-2002  
LOCUS AGENCOURT 6543030 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5548943  
DEFINITION 5', mRNA sequence.

ACCESSION BM806632

VERSION BM806632.1 GI:19123455

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
Tissue Procurement: ATCC/DCTP/DTF  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12258 row: k column: 24  
High quality sequence stop: 649.  
Location/Qualifiers  
1. 1050

FEATURES  
source  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5548943"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_72"  
/notes="Organ: skin; Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2 kb. Library constructed by Life Technologies."

## ORIGIN

Query Match 39.0%; Score 905.2; DB 4; Length 1050;  
Best Local Similarity 97.6%; Pred. No. 6.8e-221;  
Matches 961; Conservative 0; Mismatches 19; Indels 5; Gaps 4;  
QY 865 ATATCTCTGTAGAGATTACAGGTAATTTGTAGAGAGTAAGCAATTTCTCTGGGAG 924  
Db 1 ATATCTCTGTAGAGATTACAGGTAATTTGTAGAGAGTAAGCAATTTCTCTGGGAG 60

QY 925 AACAGCAGGAGTACCAACAGAGAAACAAATAGAAAAACAGATGATCCAGAACAAAAAGCAA 984  
Db 61 AACAGCAGGAGTACCAACAGAGAAACAAATAGAAAAACAGATGATCCAGAACAAAAAGCAA 120  
QY 985 AAGTTAAGAAAAAGAACCCCTAAACCTTTTAAATTAATTTGATTAAGACTATTAAAGCTGTAAC 1044  
Db 121 AAGTTAAGAAAAAGAACCCCTAAACCTTTTAAATTAATTTGATTAAGACTATTAAAGCTGTAAC 180  
QY 1045 TTGATGCTGCAGAAAAAATCCGTTAAAGGGGAAAAATTTGAGGAGCAGTGAATGCAATTTA 1104  
Db 181 TTGATGCTGCAGAAAAAATCCGTTAAAGGGGAAAAATTTGAGGAGCAGTGAATGCAATTTA 240  
QY 1105 AAGAACTTAGTACGCAAAATACCTCAGAGTCCACGAGCAAGATATGGGAAAGGGCGAGTGTG 1164  
Db 241 AAGAACTTAGTACGCAAAATACCTCAGAGTCCACGAGCAAGATATGGGAAAGGGCGAGTGTG 300  
QY 1165 AGAGTATTTGGCTGAGAAAGAGGAGAGTAATGAGGTGCTACGTTGGAGCCATCGAGACCT 1224  
Db 301 AGAGTATTTGGCTGAGAAAGAGGAGAGTAATGAGGTGCTACGTTGGAGCCATCGAGACCT 360  
QY 1225 ACCAAGAGGTGGCCAGCTTACCTGATGTCCTGAGACCTGCTGAAAGCTGAGTTTGAAGC 1284  
Db 361 ACCAAGAGGTGGCCAGCTTACCTGATGTCCTGAGACCTGCTGAAAGCTGAGTTTGAAGC 420  
QY 1285 GTGCTCAGACAGGCAACAAATTTCTAGGTGATATGAGAGGTTCCTGCTTACCTGCTGAGA 1344  
Db 421 GTGCTCAGACAGGCAACAAATTTCTAGGTGATATGAGAGGTTCCTGCTTACCTGCTGAGA 480  
QY 1345 GATTAGTTCAATTTTCCCAATGATCTCTTAAAAAATGACTTGGGGTGGGATACC 1404  
Db 481 GATTAGTTCAATTTTCCCAATGATCTCTTAAAAAATGACTTGGGGTGGGATACC 540  
QY 1405 TCTTGTAGAGAGATATGACAAATGCAAAAGAAAGTTTATGAAGAGGTGCTGAGTGTGACAC 1464  
Db 541 TCTTGTAGAGAGATATGACAAATGCAAAAGAAAGTTTATGAAGAGGTGCTGAGTGTGACAC 600  
QY 1465 CTATGATGCTTTGCTAAAGTCCATTATGCTTTCATCTGAAAGGCAAGAACAAATTTG 1524  
Db 601 CTATGATGCTTTGCTAAAGTCCATTATGCTTTCATCTGAAAGGCAAGAACAAATTTG 660  
QY 1525 CTGAGAGCATCCCATATTTAAAGAGAGATAGAAATCCGAGATCTTGGCAGCTGATGATG 1584  
Db 661 CTGAGAGCATCCCATATTTAAAGAGAGATAGAAATCCGAGATCTTGGCAGCTGATGATG 720  
QY 1585 GGAGATTTTATTTCCACCTGGGGGATGCCATGCGAGAGGTGGGAAACAAAGAGGCATATA 1644  
Db 721 GGAGATTTTATTTCCACCTGGGGGATGCCATGCGAGAGGTGGGAAACAAAGAGGCATATA 780  
QY 1645 AGTGGTATGAGCTTTGGGCACAAAGAGAGGACACTTTTGCATCTGCTGCGCAACCTCACTCT 1704  
Db 781 AGTGGTATGAGCTTTGGGCACAAAGAGAGGACACTTTTGCATCTGCTGCGCAACCTCACTCT 840  
QY 1705 ACAATGTGATGAGCTTGAAGCAGACAGCTTGGTGGACCCCAAGAGAAAGCGGCTACACAG 1764  
Db 841 ACAATGTG-ATGAGCTGAAAGCAGACCTTTGGTGGACCCCAAGAGAAAGCGGCTACACAG 899  
QY 1765 AGTTAGTAAAGT-CTTTTGAAGAAACCTGGAAGTTAAT-CCGAGATGAAGGCC--TTGCA 1820  
Db 900 AGTTAGTAAAGTCTTTTGAAGAAACCTGGAAGTTAATCCCGAAATGAAGGCCCTGTGCA 959  
QY 1821 GTGATGATAAGCCCAAGGCTCTCT 1845  
Db 960 GTGATGATAAGCCCAAGGCGCT 984

## RESULT 8

BM806632

LOCUS

DEFINITION

ACCESSION

VERSION

BM806636 Homo sapiens PLACENTA Homo sapiens cDNA clone CL08A0072C04  
S-PRIME, mRNA sequence.  
BX403636.2 GI:46924400



```

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1011)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 13, 2003 this sequence version replaced gi:30635043.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4537.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CL0BA007ZC04RPl&c=4537.r.

FEATURES
Location/Qualifiers
1..1011
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CL0BA007ZC04"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 37.6%; Score 873.4; DB 5; Length 1011;
Best Local Similarity 97.6%; Pred. No. 9.9e-213;
Matches 880; Conservative 5; Mismatches 17; Indels 0; Gaps 0;

QY 1 CGGACCGTGCATGGCCAGCGTGAAGATGCCAGAGCGCGGCAACAGCAGCAGCAGCG 60
DB 110 CGGACCGTGCATGGCCAGCGTGAAGATGCCAGAGCGCGGCAACAGCAGCAGCAGCG 169
QY 61 GCTCGGCGAGGTAGCAGAGTGGCGCAGCAGCGCCGCGCGCCGCGGAGAGACAA 120
DB 170 GCTCGGCGAGGTAGCAGAGTGGCGCAGCAGCGCCGCGCGCCGCGGAGAGACAA 229
QY 121 AGCATGGAGGACACAAGAAATGGGAGAAAGCGGACTCTCGGGAACCTTCATTTCA 180
DB 230 AGCATGGAGGACACAAGAAATGGGAGAAAGCGGACTCTCAGGAACCTTCATTTCA 289
QY 181 GGTATTATGGTATTCATTCGTGGCGCTCGACATCTGTAGCTGTGTTGGTTGATC 240
DB 290 GGTATTATGGTATTCATTCGTGGCGCTCGACATCTGTAGCTGTGTTGGTTGATC 349
QY 241 TTGTTGACTATGAGGAAGTTCTAGAAACTAGGAATCTATGATGCTGATGCTGATGG 300
DB 350 TTGTTGACTATGAGGAAGTTCTAGAAACTAGGAATCTATGATGCTGATGCTGATGG 409
QY 301 ATTTTGTATGTGATGATCCAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTCAG 360
DB 410 ATTTTGTATGTGATGATCCAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTCAG 469
QY 361 CAGCAGTCCCGCAGAGAGGTGAGCCACACACTGAGCCCGCAGGAGCAGGTTCTCTGG 420
DB 470 CAGCAGTCCCGCAGAGAGGTGAGCCACACACTGAGCCCGCAGGAGCAGGTTCTCTGG 529
QY 421 AGGCAGAAACCCAGAAATATCGAAGATGAAGCAAAAGAAACAAATTTCAGTCCCTTC 480
DB 530 AGGCAGAAACCCAGAAATATCGAAGATGAAGCAAAAGAAACAAATTTCAGTCCCTTC 589

```

```

QY 481 AAATGGTACACCGCAGAA CATGTTGAGGGAGAGACTTGCACCAAGAGATGGACCCACAG 540
DB 590 AAATGNTATCACCGCAGAA CATGTTGAGGGAGAGACTTGCACCAAGAGATGGACCCACAG 649
QY 541 GAGAACCAACAAAGAGGATGATGATGTTCTTATGGCGACTGATGTAGATGATGATTTG 600
DB 650 GAGAACCAACAAAGAGGATGATGATGTTCTTATGGCGACTGATGTAGATGATGATTTG 709
QY 601 AGACCCCTGGAACCTGAAAGTATCTCATGAAGAAACCGAGCATAGTTACCAACGTTGAAGAGA 660
DB 710 AGACCCCTGGAACCTGAAAGTATCTCATGAAGAAACCGAGCATAGTTACCAACGTTGAAGAGA 769
QY 661 CAGTTTCAACAAGACTCTATCAGGATATCGAAGAGATGATGCTGAGCAGGAAATCCAG 720
DB 770 CAGTTTCAACAAGACTCTATCAGGATATCGAAGAGATGATGCTGAGCAGGAAATCCAG 829
QY 721 ATTCCAGTGAACCAAGTAGTAGAAGATTGCACCATGATACAGATGATGTAACAT 780
DB 830 ATTCCAGTGAACCAAGTAGTAGAAGATTGCACCATGATACAGATGATGTAACAT 889
QY 781 ACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA 840
DB 890 ACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA 949
QY 841 CAGAAGTAACTGCTCCCTCAGGATATCTCTGTAAGATTCACAGGTAATTTGTAGAAG 900
DB 950 CAGAAGTAACTGCTCCCTCAGGATATCTCTGTAAGATTCACAGGTAATTTGTAGAAG 1009
QY 901 AA 902
DB 1010 TA 1011

RESULT 9
LOCUS BX459083
DEFINITION BX459083 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE012YB13
5-PRIME, mRNA sequence.
ACCESSION BX459083
VERSION BX459083.2 GI:47064816
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1085)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 22, 2003 this sequence version replaced gi:31027087.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4537.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DE012CA07QPl&c=4537.r.

FEATURES
Location/Qualifiers
1..1085
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE012YB13"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into

```



the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

## ORIGIN

```
Query Match      35.9%; Score 834.6; DB 5; Length 1085;
Best Local Similarity 99.3%; Pred. No. 9.3e-203;
Matches 837; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 113 AGAGACAAAGCATGGAGGACACAAAGATGGAGGAAAGCGGACTCTCGGAACTTCATT 172
DB 1 AGAGACAAAGCATGGAGGACACAAAGATGGAGGAAAGCGGACTCTCAGGAACTTCATT 60
QY 173 CTTACAGTGGTTTATGGTGATTCATTGCTGGGGCTCTGGACATCTGTAGCTGCTGTTG 232
DB 61 CTTACAGTGGTTTATGGTGATTCATTGCTGGGGCTCTGGACATCTGTAGCTGCTGTTG 120
QY 233 GTTTGATCTTTGACTATGAGGAAGTCTTGGGAAACTAGGAATCTATGATGCTGATGG 292
DB 121 GTTTGATCTTTGACTATGAGGAAGTCTTGGGAAACTAGGAATCTATGATGCTGATGG 180
QY 293 TGATGGAGATTTGATGATGATGATGCCAAAGTTTTTATAGGACTTAAAGAGAGATCTAC 352
DB 181 TGATGGAGATTTGATGATGATGATGCCAAAGTTTTTATAGGACTTAAAGAGAGATCTAC 240
QY 353 TTCAGACCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGT 412
DB 241 TTCAGACCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGT 300
QY 413 TCCTGTGGAGGAGAACCCAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCT 472
DB 301 TCCTGTGGAGGAGAACCCAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCT 360
QY 473 TCTCATGAATGTTACGCGAGACATGTTGAGGAGAGACTTGCACAGAGATGG 532
DB 361 TCTCATGAATGTTACGCGAGACATGTTGAGGAGAGACTTGCACAGAGATGG 420
QY 533 ACCACAGGAGAACCCAAAGAGATGATGAGTTTCTTATGGCGACTGATGATGA 592
DB 421 ACCACAGGAGAACCCAAAGAGATGATGAGTTTCTTATGGCGACTGATGATGA 480
QY 593 TAGATTTGAGACCTCGAACTGAAATATCTCATGAAGAAACCGAGCATAGTTACCACT 652
DB 481 TAGATTTGAGACCTCGAACTGAAATATCTCATGAAGAAACCGAGCATAGTTACCACT 540
QY 653 GGAAGACAGAGTTTCAAGACTGTAATCAGGATATGGAAGAGATGATGTTGAGCAGA 712
DB 541 GGAAGACAGAGTTTCAAGACTGTAATCAGGATATGGAAGAGATGATGTTGAGCAGA 600
QY 713 AAATCCAGATTCAGTGAACCAAGTAGTAGAGATGAAGATTTGCCACCATGATACAGATGA 772
DB 601 AAATCCAGATTCAGTGAACCAAGTAGTAGAGATGAAGATTTGCCACCATGATACAGATGA 660
QY 773 TGTAACATACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAAAATGAAGGGAT 832
DB 661 TGTAACATACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAAAATGAAGGGAT 720
QY 833 AGAAATCACAAGAGTAACTGCTCCCTGAGGATATCTTGTAGAGATTTCACAGTAAAT 892
DB 721 AGAAATCACAAGAGTAACTGCTCCCTGAGGATATCTTGTAGAGATTTCACAGTAAAT 780
QY 893 TGTAAGAGAGTAACTGCTCCCTGAGGATATCTTGTAGAGATTTCACAGTAAAT 952
DB 781 TGTAAGAGAGTAACTGCTCCCTGAGGATATCTTGTAGAGATTTCACAGTAAAT 840
QY 953 TAG 955
DB 841 AAG 843
```

## RESULT 10

AUI30952

LOCUS

DEFINITION AUI30952 NT2RP3 Homo sapiens cDNA clone NT2RP3001731 5', mRNA

```
sequence.
AUI30952
AUI30952.1 GI:10991306
EST.
Homo sapiens (human)
Homo sapiens
Mammalia; Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,
Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and
Isogai, T.
HRI human cDNA project
Unpublished (2000)
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
Location/Qualifiers
1..859
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xrefs="taxon:9606"
/clone="NT2RP3001731"
/cell_type="teratocarcinoma"
/clone_lib="NT2RP3"
/notes="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
cells after 2-weeks retinoic acid (RA) induction"

FEATURES
Source
Query Match 33.9%; Score 787.6; DB 1; Length 859;
Best Local Similarity 96.5%; Pred. No. 1e-190;
Matches 834; Conservative 0; Mismatches 25; Indels 5; Gaps 3;

QY 1013 AAATAAATTTGATAAGACTATTAAAGCTGAACCTTGACTGCGCAAAAACCTCCGTAAAG 1072
DB 1 AAATAAATTTGATAAGACTATTAAAGCTGAACCTTGACTGCGCAAAAACCTCCGTAAAG 60
QY 1073 GGGAAAAATTTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGCGCAAAATACCTCAGAG 1132
DB 61 GGGAAAAATTTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGCGCAAAATACCTCAGAG 120
QY 1133 TCCACGAGCAAGATATGGGAAGCGCAGTGTGAGGATGATTGGCTGAGAAGAGGAGAAG 1192
DB 121 TCCACGAGCAAGATATGGGAAGCGCAGTGTGAGGATGATTGGCTGAGAAGAGGAGAAG 180
QY 1193 TAATGAGTGCTACGTGGAGCCATCGAGACCTACCAAGAGTGGCCAGCTACCTGATGT 1252
DB 181 TAATGAGTGCTACGTGGAGCCATCGAGACCTACCAAGAGTGGCCAGCTACCTGATGT 240
QY 1253 CCCTGCGAGCTCTGAACCTGAGTTTGAACGCTCGCTCAGACGGCAACAATTTCTAGG 1312
DB 241 CCCTGCGAGCTCTGAACCTGAGTTTGAACGCTCGCTCAGACGGCAACAATTTCTAGG 300
QY 1313 TCATATGAGAGGTTCCCTGCTTACCTCGCAGAGATTAGTTCAACTATTTTCCCAATGATAC 1372
DB 301 TCATATGAGAGGTTCCCTGCTTACCTCGCAGAGATTAGTTCAACTATTTTCCCAATGATAC 360
QY 1373 TTCCTTAAAAAATGACCTTGGCGTGGGATACCTCTGATAGGAGATAATGACAATGCCAAA 1432
DB 361 TTCCTTAAAAAATGACCTTGGCGTGGGATACCTCTGATAGGAGATAATGACAATGCCAAA 420
QY 1433 GAAAGTTTATGAGAGGCTGCTGAGTGTGACACCTTAATGATGGCTTTTCTTAAGTCCATTA 1492
DB 421 GAAAGTTTATGAGAGGCTGCTGAGTGTGACACCTTAATGATGGCTTTTCTTAAGTCCATTA 480
```

```
Qy 1493 TGGCTTCATCTGAGGACAGACAAAATGCTGAGAGCATCCCATATTTAAAGGAAGG 1552
Db 481 TGGCTTCATCTGAGGACAGACAAAATGCTGAGAGCATCCCATATTTAAAGGAAGG 540
Qy 1553 AATAGAATCCGGAGATCCTGGCACTGATGATGGAGAGATTTATTTCCACCTGGGGGATGC 1612
Db 541 AATAGAATCCGGAGATCCTGGCACTGATGATGGAGAGATTTATTTCCACCTGGGGGATGC 600
Qy 1613 CATCAGAGGTTGGGACAAAGAGGATATAAGTGTATGAGCTGGGACAGAGAGG 1672
Db 601 CATCAGAGGTTGGGACAAAGAGGATATAAGTGTATGAGCTGGGACAGAGAGG 660
Qy 1673 ACATTTTGATCTCTCTGGCAACCTCACTCTACAATGTGAATGGAATGAAAGCACAGCC 1732
Db 661 ACATTTTGATCTCTCTGGCAACCTCACTCTACAATGTGAATGGAATGAAAGCACANCC 720
Qy 1733 TTGTGTGACCCCAAAAGAAAGCGGCTACACAGATTTAGTAAAGTCTTTAGAAAGAACTG 1792
Db 721 TTGTGTGACCCCAAAAGAAAGCGGCTACACAGATTTAGTAAAGTCTTTAGAAAGAACTG 779
Qy 1793 GAAGTTAATCCAGATGAAGGCTTGAGTATGATGATGAATGAAGGCTCTTCTCTGCC 1852
Db 780 GAAGTTAATCCAGATGAAGGCTTGAGTATGATGATGAATGAAGGCTCTTCTCTGCC 835
Qy 1853 TGAGATGAACCTGAGGGGAAAA 1876
Db 836 TGAGATGAACCTGAGGGGAAAA 859

RESULT 11
BU838423
LOCUS
DEFINITION BU838423 930 bp mRNA linear EST 16-OCT-2002
clone IMAGE:6177038 5', mRNA sequence.
ACCESSION BU838423
VERSION BU838423.1 GI:24022818
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 930)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM13554 row: f column: 15
High quality sequence start: 66
High quality sequence stop: 549.
Location/Qualifiers
1..930
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6177038"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupski dorsal root ganglion"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
Not1; Site_2: Salt; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGCTCCG-3' and
```

```
5'-GACTAGTCTTAGATCCGAGCGGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
ORIGIN
Query Match 31.0%; Score 719.8; DB 5; Length 930;
Best Local Similarity 96.3%; Pred. No. 2.7e-173;
Matches 758; Conservative 0; Mismatches 27; Indels 2; Gaps 2;
Qy 1 CGACCGTGCATATGGCCGCGTAAGATGCGCAAGACGCGCCACAGCAGCAGCAGCG 60
Db 138 CGACCGTGCATATGGCCGCGTAAGATGCGCAAGACGCGCCACAGCAGCAGCAGCG 197
Qy 61 GCTCCGCGAGCGGTAGCAGAGTGGCGGACAGCAGCCCGCGGCGCCGAGAGAGACAA 120
Db 198 GCTCCGCGAGCGGTAGCAGAGTGGCGGACAGCAGCCCGCGGCGCCGAGAGAGACAA 257
Qy 121 AGCATGGAGGACACAAGAATGGAGGAAAGCGGACTCTCGGGAACCTTCATTTCACGT 180
Db 258 AGCATGGAGGACACAAGAATGGAGGAAAGCGGACTCTCAGGAACCTTCATTTCACGT 317
Qy 181 GGTATTATGATGATGATGCTGGCGGTCTGGACATCTGTAGCTGCTGTTGTTGATC 240
Db 318 GGTATTATGATGATGATGCTGGCGGTCTGGACATCTGTAGCTGCTGTTGTTGATC 377
Qy 241 TTGTTGACTATGAGGAAGTCTAGGAAACTAGGAATCTATGATGCTGATGTTGATGGAG 300
Db 378 TTGTTGACTATGAGGAAGTCTAGGAAACTAGGAATCTATGATGCTGATGTTGATGGAG 437
Qy 301 ATTTTGATGTGGATGATGCCAAAGTTTTATTTAGGACTTTAAAGAGAGATCTACTTCAGAGC 360
Db 438 ATTTTGATGTGGATGATGCCAAAGTTTTATTTAGGACTTTAAAGAGAGATCTACTTCAGAGC 497
Qy 361 CAGCAGTCCCGCAGAAAGAGGCTGACCAACACTGAGCCCGGAGGAGCAGGTTCTGTGG 420
Db 498 CAGCAGTCCCGCAGAAAGAGGCTGACCAACACTGAGCCCGGAGGAGCAGGTTCTGTGG 557
Qy 421 AGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480
Db 558 AGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 617
Qy 481 AATATGTTACGCGAGAACACTGTTGAGGAGAGAGACTTGCACCAAGAGATGGA-CCACA 539
Db 618 AATATGTTACGCGAGAACACTGTTGAGGAGAGAGACTTGCACCAAGAGATGGA-CCACA 677
Qy 540 GGAGAACCAACACAGAGGATGATGATTTCTTATGCGGACTGATGTAGATGATAGATTT 599
Db 678 GGAGAACCAACACAGAGGATGATGATTTCTTATGCGGACTGATGTAGATGATAGATTT 737
Qy 600 GAGACCTTGGAACTGTAAGTATCTCATGAAGAACCCGAGCATAGTTACCACTGGAAAGAG 659
Db 738 GAGACCTTGGCACTTGAAGTATCTCATGAAGAACCCGAGCATAGTTACCACTGGAAAGAG 797
Qy 660 ACAGTTTTCACAGACTGTAATCAGGATATCGAAGAGATGATGTGTAGAGAGGAAATCCA 719
Db 798 ACAGTTTTCACAGACTGTAATCAGGATATCGAAGAGATGATGTGTAGAGAGGAAATCCA 857
Qy 720 GATTCCAGTGAACCAAGTAGTAGAGATG-AAAAGTTGCACCATGATACAGATGATGAAC 778
Db 858 GATTCCAGTGAACCAAGTAGTAGAGATG-AAAAGTTGCACCATGATACAGATGATGAAC 917
Qy 779 ATACCAA 785
Db 918 ATATCCA 924

RESULT 12
CK724966/c
LOCUS
DEFINITION UI-H-BD1-axp-c-22-0-UI.s1 NCI_CGAP_ED1 Homo sapiens cDNA clone
```

```

ACCESSION      CK724966
VERSION         CK724966.1
KEYWORDS        GI:42535838
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE       1 (bases 1 to 782)
AUTHORS         NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE           National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
                Tumor Gene Index
JOURNAL         Unpublished (1997)
COMMENT         Contact: Daniela S. Gerhard, Ph.D.
                Office of Cancer Genomics
                National Cancer Institute / NIH
                Bldg: 31 Rm10A07 Bethesda, MD 20892
                Email: cgapbs-re@mail.nih.gov
                Tissue Procurement: Dr. Jose Mercuende
                cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
                cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
                DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
                Clone Distribution: Distribution information can be found at
                http://genome.uiowa.edu/distribution/cgap.html
                Seq primer: M13 FORWARD
                POLYA=Yes.

FEATURES             Location/Qualifiers
     source           1..782
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-EDI-axp-c-22-0-UI"
                     /tissue_type="Chondrosarcoma"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI CGAP EDI"
                     /note="Organ: Left Pubic Bone; Vector: pVT73-Pac
                     (Pharmacia) with a modified polylinker; Site 1: EcoR I;
                     Site 2: Not I; NCI CGAP EDI is a normalized cDNA library
                     containing the following tissue(s): Chondrosarcoma cell
                     line C85. The library was constructed according to
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. First strand cDNA synthesis was primed with an
                     oligo-dT primer containing a Not I site. Double stranded
                     cDNA was ligated to an EcoR I adaptor, digested with Not
                     I, and cloned directionally into pVT73-pac vector. The
                     oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is GCTCAGGCT.
                     TAG TISSUE=chondrosarcoma
                     TAG LIB=UI-H-EDI
                     TAG_SEQ=CCTCAGGCT"

ORIGIN
Query Match      30.0%; Score 697; DB 7; Length 782;
Best Local Similarity 98.9%; Pred. No. 1.9e-167;
Matches 709; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 1210 GAGCCATCGAGACTACCAAGAGGTGGCCAGCTTACCTGATGTCCTCGAGACCTGCTGA 1269
      |||||
DB 716 GAGCCATCGAGACNTACCAAGA-GTGGCCAGCTTACCTGATGTCCTCGAGACCTGCTGA 658
      |||||
QY 1270 AGCTGAGTTTGAAGCTCGCTCAGACAGGCAACAATTTCTAGTCTATATGAGGTTCCC 1329
      |||||
DB 657 AGCTGAGTTTGAAGCTCGCTCAGACAGGCAACAATTTTAGTCTATATGAGGTTCCC 598
      |||||
QY 1330 TGCTTACCCTCGAGAGTAGTTCACTATTTCCTTCCCAATGATATCTTCTTAAAAAATGACC 1389
      |||||
DB 597 TGCTTACCCTCGAGAGTAGTTCACTATTTCCTTCCCAATGATATCTTCTTAAAAAATGACC 538
      |||||
QY 1390 TTGCGGTGGGATACCTCTTGATAGGAGATAATGCAATGCAAGAAAGTTTATCAGAGG 1449
      |||||

```

```

537 TTGGCGTGGGATACCTCTTGATAGGAGATAATGCAATGCAAGAAAGTTTATGAAGAGG 478
1450 TGCTGAGTGTGACACCTTAATGATGGCTTTGCTTAAAGTCCATTTATGGCTTTCATCTGAAGG 1509
477 TGCTGAGTGTGACACCTTAATGATGGCTTTGCTTAAAGTCCATTTATGGCTTTCATCTGAAGG 418
1510 CACGAACAAAATTTGCTGAGAGCATCCCATATTTTAAAGGAAGGAATAGAAATCCGAGATC 1569
417 CACGAACAAAATTTGCTGAGAGCATCCCATATTTTAAAGGAAGGAATAGAAATCCGAGATC 358
1570 CTGGCACTGATGATGGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGAGGTTGGGA 1629
357 CTGGCACTGATGATGGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGAGGTTGGGA 298
1630 ACAAGAGGCATATAAGTGTATGAGCTTGGGCHCAGAGAGGACACTTTTGCATCTGTCT 1689
297 ACAAGAGGCATATAAGTGTATGAGCTTGGGCHCAGAGAGGACACTTTTGCATCTGTCT 238
1690 GGCACCTCTCACTCTCAATGTAATGCACTGAAAGCACAGCCCTTGTGGACCCCAAAAG 1749
237 GGCACCTCTCACTCTCAATGTAATGCACTGAAAGCACAGCCCTTGTGGACCCCAAAAG 178
1750 AAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAAACTGGAAGTTAAATCCGAGATG 1809
177 AAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAAACTGGAAGTTAAATCCGAGATG 118
1810 AAGGCTTGCAGTGATGATAAGCCAAAGTCTCTTCTGCTGAGGATGAACCTGA 1869
117 AAGGCTTGCAGTGATGATAAGCCAAAGTCTCTTCTGCTGAGGATGAACCTGA 58
1870 GGGAAAAGGGGACTGGAGCCAGTTACGCTGTGGCAGCAAGGAAGAAATGAAA 1926
57 GGGAAAAGGGGACTGGAGCCAGTTACGCTGTGGCAGCAAGGAAGAAATGAAA 1

```

```

RESULT 13
BM790814      689 bp mRNA linear EST 05-MAR-2002
LOCUS         K-EST0070696 S21SNUS20 Homo sapiens cDNA clone S21SNUS20-16-C08 5',
DEFINITION    mRNA sequence.
ACCESSION     BM790814
VERSION        BM790814.1 GI:19139046
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 689)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
JOURNAL        Unpublished (2002)
COMMENT        Contact: Kim YS
                Genome Research Center
                Korea Research Institute of Bioscience & Biotechnology
                52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
                Tel: +82-42-860-4470
                Fax: +82-42-860-4409
                Email: yongsung@mail.kribb.re.kr
                Plate: 16 row: C column: 08
                High quality sequence stop: 689.
                Location/Qualifiers
                1..689
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="S21SNUS20-16-C08"
                /sex="P"
                /tissue_type="Stomach"
                /cell_type="Floating aggregates"
                /cell_line="SNU-520"
                /lab_host="Top10r"

```

/clone\_lib="s21SNUS20"  
/note="Organ: Stomach; Vector: pTZ18RP1; Site\_1: EcoRI;  
Site\_2: NotI; The poly (A)+ RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then decapped  
with cabacco acid pyrophosphatase (TAP). The decapped  
intact mRNA was ligated with DNA-RNA linker including EcoR  
I site by treatment of T4 RNA ligase and the first strand  
cDNA was synthesized from oligo dt-selected mRNA by  
priming with dt-tailed vector. The dt-tailed vector was  
adjusted to have about 60nt. The cDNA vector was  
circularized with E. coli DNA ligase after digestion of  
EcoRI which site is also included in vector. An RNA strand  
converted to a DNA strand by Okayama-Berg method. The  
obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10<sup>+</sup> by electroporation method.  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library."

ORIGIN

Query Match 29.6%; Score 689; DB 4; Length 689;  
Best Local Similarity 100.0%; Pred. No. 2e-165;  
Matches 689; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1219 AGACCTACCAAGAGTGGCCAGCTTACTGTATGTCCTCGACACCTCTGCTGAAGCTGAGTT 1278  
DB 1 AGACCTACCAAGAGTGGCCAGCTTACTGTATGTCCTCGACACCTCTGCTGAAGCTGAGTT 60  
QY 1279 TGAAGCGTCGCTCAGACAGGCAACAAATTTAGTGTCATATGAGAGGTTCCCTGCTTACCC 1338  
DB 61 TGAAGCGTCGCTCAGACAGGCAACAAATTTAGTGTCATATGAGAGGTTCCCTGCTTACCC 120  
QY 1339 TGCAGAGATTAGTTCAACTATTTCCCAATGATATCTCTTAAATAAAGTACCTTGGCGTGG 1398  
DB 121 TGCAGAGATTAGTTCAACTATTTCCCAATGATATCTCTTAAATAAAGTACCTTGGCGTGG 180  
QY 1399 GATACCTCTTGATAGAGATAATGACATGCAATGCAAAAGAAATTTATGAAGAGTGCTGAGTG 1458  
DB 181 GATACCTCTTGATAGAGATAATGACATGCAATGCAAAAGAAATTTATGAAGAGTGCTGAGTG 240  
QY 1459 TGACACCTTAATGATGGCTTGTAAAGTCCATTATGCTTCCATCTGAGGCAAGCAAGACA 1518  
DB 241 TGACACCTTAATGATGGCTTGTAAAGTCCATTATGCTTCCATCTGAGGCAAGCAAGACA 300  
QY 1519 AAATTTGCTGAGAGATCCCATATTTAAAGGAAGAAATAGAAATCCGGAGATCCTGGCACTG 1578  
DB 301 AAATTTGCTGAGAGATCCCATATTTAAAGGAAGAAATAGAAATCCGGAGATCCTGGCACTG 360  
QY 1579 ATGATGGGAGATTTTATTTCCACCTGGGGGATGCCATGAGAGGTTGGGAACAAAGAGG 1638  
DB 361 ATGATGGGAGATTTTATTTCCACCTGGGGGATGCCATGAGAGGTTGGGAACAAAGAGG 420  
QY 1639 CATATAAGTGGTATGAGCTTGGGCAAGAGAGGACATTTGTCATCTCTGTCGGCAAGCCT 1698  
DB 421 CATATAAGTGGTATGAGCTTGGGCAAGAGAGGACATTTGTCATCTCTGTCGGCAAGCCT 480  
QY 1699 CACTCTACAAATGTGAATGACTGAAAGACACAGCTTGTGGACCCCAAAAGAAACGGGCT 1758  
DB 481 CACTCTACAAATGTGAATGACTGAAAGACACAGCTTGTGGACCCCAAAAGAAACGGGCT 540  
QY 1759 ACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAATCCGAGATGAGGCTTTG 1818  
DB 541 ACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAATCCGAGATGAGGCTTTG 600  
QY 1819 CAGTGATGTAAGCAAGGTCCTCTGCTGAGGATGAAACCTGAGGGAAGAAAG 1878  
DB 601 CAGTGATGTAAGCAAGGTCCTCTGCTGAGGATGAAACCTGAGGGAAGAAAG 660  
QY 1879 GGGACTGGAGCCAGTTTCAGCTGTGGCAG 1907  
DB 661 GGGACTGGAGCCAGTTTCAGCTGTGGCAG 689

BUI66117  
LOCUS BUI66117 918 bp mRNA linear EST 04-SEP-2002  
DEFINITION AGENCOURT 7914661 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:6147375  
5' mRNA sequence.  
BUI66117  
VERSION BUI66117.1 GI:22680069  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 918)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
Tissue Procurement: ATCC/DCTD/DTP  
cDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM13477 row: b column: 16  
High quality sequence stop: 634.  
Location/Qualifiers  
1..918  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6147375"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 72"  
/note="Organ: skin; Vector: pCW-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 2 kb. Library constructed by Life  
Technologies."  
FEATURES  
Source  
1  
Query Match 28.7%; Score 667.8; DB 5; Length 918;  
Best Local Similarity 91.3%; Pred. No. 6.2e-160;  
Matches 775; Conservative 0; Mismatches 14; Indels 60; Gaps 4;  
QY 1 CGGACCGTGCATGGCCAGCGTAAGATGCCAAGAGCAGCGCCGCGGCGCGGAGACAGCAGCAGCG 60  
DB 118 CGGACCGTGCATGGCCAGCGTAAGATGCCAAGAGCAGCGCGCGGCGCGGAGACAGCAGCAGCG 177  
QY 61 GCTCCGCGCAGCGGTAGCAGCGTGGCGGAGCGAGCGCGCGCGCGCGGCGCGGAGACAGCAGCAG 120  
DB 178 GCTCCGCGCAGCGGTAGCAGCGTGGCGGAGCGAGCGCGCGCGCGGCGCGGAGACAGCAGCAG 237  
QY 121 AGCATGAGGACACAAAGATGGGAGGAAAGCGGAGCTCTCGGGAACCTTCACTTTCACGT 180  
DB 238 AGCATGAGGACACAAAGATGGGAGGAAAGCGGAGCTCTCGGGAACCTTCACTTTCACGT 237  
QY 181 GGTATTATGTTGATGTCATCTGCTGGCGCTCTGGGACATCTAGCTGTGCTGTTGTTGTTGATC 240  
DB 298 GGTATTATGTTGATGTCATCTGCTGGCGCTCTGGGACATCTAGCTGTGCTGTTGTTGTTGATC 357  
QY 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGCTGATGGAG 300  
DB 358 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGCTGATGGAG 417  
QY 301 ATTTTGTGATGATGATGTCCTCAAGATTTTATAGGACTTAAAGAGAGATCTACTTTCAGAGC 360  
DB 418 ATTTTGTGATGATGATGTCCTCAAGATTTTATAGGACTTAAAGAGAGATCTACTTTCAGAGC 477  
QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCGCGGAGCAGGCTTCTCTGAG 420  
DB 478 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCGCGGAGCAGGCTTCTCTGAG 537



Qy	944	AGAAACAAATAGAAAAACAGATGATCCAGAACAAAAAGCAAAAGTTAAGAAAAAGAGCC	1003
Db	552	AGAAACAAATAGAAAAACAGATGATCCAGAACAAAAAGCAAAAGTTAAGAAAAAGAGCC	611
Qy	1004	TAAACTTTTAAATAAATTTGATAAGACTATTAAAGCTGAACTTGATGCTGCAGAAAAACT	1063
Db	612	TAAACTTTTAAATAAATTTGATAAGACTATTAAAGCTGAACTTGATGCTGCAG-AAAACT	670
Qy	1064	CCGTAAAGGGGCAAAATTTGAGGAGCAGTCAAT	1097
Db	671	CCGTAAAGGGGCAAAATTTGAGGAGCAGTCAAT	704

Search completed: March 25, 2005, 14:28:59  
Job time : 6791 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 24, 2005, 12:41:52 ; Search time 1157 Seconds  
(without alignments)  
11890.647 Million cell updates/sec

Title: US-09-436-184-3

Perfect score: 2324

Sequence: 1 cggaccgtgcaatgcccg.....ctgggaactctggagaga 2324

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseqn16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001s:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2324	100.0	2324	4	Aaf89811 cDNA enco
2	2324	100.0	2324	8	ABs57868 Human cDN
3	2324	100.0	2324	9	ADA00640 cDNA enco
4	2320.8	99.9	2452	13	ADP23914 PRO polyp
5	2320.8	99.9	2452	13	ADP23914 Human ASP
6	2319.2	99.8	5358	10	ADJ56257 Bovine CD
7	947	40.7	2680	6	ABX04178 Human MRN
8	947	40.7	2680	13	ADP24208 PRO polyp
9	819	35.2	2442	2	Aaz23609 Human lab
10	819	35.2	2442	6	ABK39743 cDNA enco
11	819	35.2	2442	8	ACA12072 Human lun
12	819	35.2	2442	8	ACA03258 Lung canc
13	819	35.2	2442	10	ADH47310 Human lun
14	819	35.2	2442	13	ADJ21229 Human lun
15	711.4	30.6	2648	10	ACA92448 Human lun
16	617.6	26.6	3110	11	ACN88788 Breast ca
17	464.6	20.0	502	9	ACH47067 Human inf
18	416.2	17.9	660	6	ABQ59507 Human col
19	317	13.6	495	11	ADT94758 Colon can
20	213.6	9.2	1662	4	ABL10339 Drosophil

21	170.6	7.3	1382	2	AAX22124	Human sec
22	170.6	7.3	1382	8	ADA39640	Human sec
23	170.6	7.3	1382	8	ACC50349	Human sec
24	170.6	7.3	1382	10	ADA55835	Gene enco
25	153.2	6.6	161	12	ACH93769	Human gen
26	153.2	6.6	592	12	ACH80069	Human gen
27	147.6	6.4	3986	4	ABL10338	Human col
28	127.4	5.5	583	6	ABQ59571	Human col
29	113.4	4.9	1686	2	AAT05868	DNA enco
30	112.2	4.8	3399	2	AAT05868	Chicken l
31	103.2	4.4	3755	4	ABL10352	Drosophil
32	96.8	4.2	671	3	AAC95033	Cat flea
33	87.8	3.8	3579	3	AAQ70099	Plasmodiu
34	81.8	3.5	4590	1	AAAG60472	Sequence
35	74.8	3.2	1113	12	ADO00111	Novel hum
36	74.6	3.2	3095	2	ADN98542	Sequence
37	74.6	3.2	17137	6	AAQ03875	Sequence
38	72.2	3.1	31124	11	ACN44156	Human imm
39	71.2	3.1	3912	8	ACA22058	Mouse gen
40	70.4	3.0	15507	8	AAL53680	Prokaryot
41	70.2	3.0	1997	13	ADR07778	Genomic D
42	70.2	3.0	585	2	AAQ87589	Full leng
43	68.6	3.0	585	2	AAQ87589	DNA enco
44	68.6	3.0	996	2	AAQ87588	DNA enco
45	67.4	2.9	3337	2	AAT34620	P. vivax

## ALIGNMENTS

## RESULT 1

Aaf89811  
ID AAF89811 standard; cDNA; 2324 BP.

AC AAF89811;

DT 23-JUL-2001 (first entry)

XX cDNA encoding human aspartyl (asparaginyl) beta-hydroxylase (HAAH).

DE Epidermal growth factor-like domain; EGF-like domain; cancer;

KW Human aspartyl beta-hydroxylase; HAAH; malignant neoplasm; tumour; ss.

OS Homo sapiens.

PH Key

FT CDS

FT Location/Qualifiers

FT 12..2289

FT /\*tag= a

FT /product= "human aspartyl (asparaginyl) beta-hydroxylase"

XX WO200135102-A2.

XX 17-MAY-2001.

XX PD

XX PF

XX 08-NOV-2000; 2000WO-US030738.

XX PR

XX 08-NOV-1999; 99US-00436184.

XX PA (RHOD-) RHODE ISLAND HOSPITAL LIFESPAN PARTNER.

XX XX

XX Wands JR, De La Monte SM, Ince N, Carlson RI;

XX WPT; 2001-329171/34.

XX DR P-FSDS; AAB83919.

XX XX

XX Diagnosing malignant neoplasm in a mammal, involves contacting mammalian

XX sample with antibody that binds to human aspartyl beta-hydroxylase

XX polypeptide to form antigen-antibody complex and detecting the complex.

XX PS Disclosure; Page 6-7; 76pp; English.

XX The present sequence encodes a human aspartyl (asparaginyl) beta-

XX hydroxylase (HAAH) enzyme. Epidermal growth factor (EGF)-like domains of

XX CC



CC polypeptides are hydroxylated by HAAH enzymes. HAAH is used in the method  
CC of the invention. The specification describes a method for diagnosing a  
CC malignant neoplasm in a mammal. The method comprises contacting a body  
CC fluid with an antibody which binds to HAAH polypeptide under complex  
CC forming conditions, and detecting the antigen-antibody complex. The  
CC method is useful for diagnosing and prognosing a malignant neoplasm in a  
CC bodily fluid e.g. central nervous system (CNS)-derived body fluid, blood,  
CC serum, urine, saliva, sputum, lung effusion, and ascites fluid of mammal,  
CC where the neoplasm is derived from endodermal tissue and is selected from  
CC colon cancer, breast cancer, pancreatic cancer, liver cancer, cancer of  
CC bile ducts, and cancer of the CNS. HAAH antibodies, linked to a cytotoxic  
CC agent, are useful for killing tumour cells

XX  
SQ Sequence 2324 BP; 753 A; 463 C; 625 G; 483 T; 0 U; 0 Other;

Query Match 100.0%; Score 2324; DB 4; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGACCTGCAATGGCCAGCGTAAGAAATGCCAAGACAGCGGCAACAGCAGCAGCAGCG 60  
DB 1 CGGACCTGCAATGGCCAGCGTAAGAAATGCCAAGACAGCGGCAACAGCAGCAGCAGCG 60  
QY 61 GCTCCGGCAGCGGTAGCACAGAGTCGGGGCAGCAGCAGCCCCGGGGCCCGGAGAGACAA 120  
DB 61 GCTCCGGCAGCGGTAGCACAGAGTCGGGGCAGCAGCAGCCCCGGGGCCCGGAGAGACAA 120  
QY 121 AGCATGGAGGACACAAGAAATGGGAGGAAAGCGGACCTCTCGGNACTTCATCTTCACGT 180  
DB 121 AGCATGGAGGACACAAGAAATGGGAGGAAAGCGGACCTCTCGGNACTTCATCTTCACGT 180  
QY 181 GGTATTGCTGATTCGATTCGTCGGCGTCGGACATCTGTAGCTGTCTTGGTTTGATC 240  
DB 181 GGTATTGCTGATTCGATTCGTCGGCGTCGGACATCTGTAGCTGTCTTGGTTTGATC 240  
QY 241 TTGTTGACTATGAGGAAGTTCCTAGGAAATCTATGATGCTGTGATGGTGAAGGAG 300  
DB 241 TTGTTGACTATGAGGAAGTTCCTAGGAAATCTATGATGCTGTGATGGTGAAGGAG 300  
QY 301 ATTTTGTATGGATGATGCCAAAGTTTATTTAGACTTAAGAGAGATCTACTTCAGAGC 360  
DB 301 ATTTTGTATGGATGATGCCAAAGTTTATTTAGACTTAAGAGAGATCTACTTCAGAGC 360  
QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCCGAGGAGCAGGTTCCCTGG 420  
DB 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCCGAGGAGCAGGTTCCCTGG 420  
QY 421 AGGCAGAACCCAGAAATTCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCCATG 480  
DB 421 AGGCAGAACCCAGAAATTCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCCATG 480  
QY 481 AAATGGTACAGCAGAAATGTTGAGGAGAGAGCTTGCACAAAGAGATGGACCCACAG 540  
DB 481 AAATGGTACAGCAGAAATGTTGAGGAGAGAGCTTGCACAAAGAGATGGACCCACAG 540  
QY 541 GAGAACCAACAAGAGATGATGAGTTCTTATGGCGACTGATGATGATGATGATG 600  
DB 541 GAGAACCAACAAGAGATGATGAGTTCTTATGGCGACTGATGATGATGATGATG 600  
QY 601 AGACCTTGGAACTGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTGGAAGAGA 660  
DB 601 AGACCTTGGAACTGAAGTATCTCATGAAGAAACCGAGCATAGTTTACCACGTGGAAGAGA 660  
QY 661 CAGTTTCACAGACTGTAAATCAGGATATGGAAGATGATGTCGAGCAGGAAATCCAG 720  
DB 661 CAGTTTCACAGACTGTAAATCAGGATATGGAAGATGATGTCGAGCAGGAAATCCAG 720  
QY 721 ATTCCAGTGAACAGTAGTAGAAGATGAAAGATTCACCATGATACAGATGATGTAACAT 780  
DB 721 ATTCCAGTGAACAGTAGTAGAAGATGAAAGATTCACCATGATACAGATGATGTAACAT 780  
QY 781 ACCAAGTCTATGAGGAAACAAGCAGGTATATGAACTCTTAGAAATGAAGGATAGAAATCA 840  
DB 781 ACCAAGTCTATGAGGAAACAAGCAGGTATATGAACTCTTAGAAATGAAGGATAGAAATCA 840

DB 781 ACCAAGTCTATGAGGAAACAAGCAGGTATATGAACTCTTAGAAATGAAGGATAGAAATCA 840  
QY 841 CAGAAAGTAACCTGCTCCCTCGAGGATAATCTCTGTAGAAGATTTCACAGTAAATTTGAGAG 900  
DB 841 CAGAAAGTAACCTGCTCCCTCGAGGATAATCTCTGTAGAAGATTTCACAGTAAATTTGAGAG 900  
QY 901 AAGTAAGCAATTTTCTGTGGAAGAACAGCAGGAATACCAACAGAAACAATAGAAA 960  
DB 901 AAGTAAGCAATTTTCTGTGGAAGAACAGCAGGAATACCAACAGAAACAATAGAAA 960  
QY 961 CAGATGATCCGAAACAAAAAGTAAAGAAAGAGAGCTAACTTTTAAATAAAT 1020  
DB 961 CAGATGATCCGAAACAAAAAGTAAAGAAAGAGAGCTAACTTTTAAATAAAT 1020  
QY 1021 TTGATAAGACTATTAAAGCTGAACTTGATGCTGCAGAAAAAACTCCGTAAAGGGGAAAA 1080  
DB 1021 TTGATAAGACTATTAAAGCTGAACTTGATGCTGCAGAAAAAACTCCGTAAAGGGGAAAA 1080  
QY 1081 TTGAGGAACAGTGAATGCAATTTAAAGAACTAGTAGGCAATACCTCAGAGTCCACGAG 1140  
DB 1081 TTGAGGAACAGTGAATGCAATTTAAAGAACTAGTAGGCAATACCTCAGAGTCCACGAG 1140  
QY 1141 CAGATATCGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAAATTAATGAG 1200  
DB 1141 CAGATATCGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAAATTAATGAG 1200  
QY 1201 TGCTAGTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCTACCTGATGCTCCTGAG 1260  
DB 1201 TGCTAGTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCTACCTGATGCTCCTGAG 1260  
QY 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTCGCTCAGACAGGCAACAATTTCTAGGTCATATCA 1320  
DB 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTCGCTCAGACAGGCAACAATTTCTAGGTCATATCA 1320  
QY 1321 GAGGTTCCCTGCTTACCTGCAGAGATTAGTTCAACTATTTCCCAATGATACTTCTCTAA 1380  
DB 1321 GAGGTTCCCTGCTTACCTGCAGAGATTAGTTCAACTATTTCCCAATGATACTTCTCTAA 1380  
QY 1381 AAAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATTAATGCAATGCAAGAAAGTTT 1440  
DB 1381 AAAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATTAATGCAATGCAAGAAAGTTT 1440  
QY 1441 ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGCGCTTGTCTAAAGTCCATTATGGCTTCA 1500  
DB 1441 ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGCGCTTGTCTAAAGTCCATTATGGCTTCA 1500  
QY 1501 TCCTGAAGGACAGAAACAAATTTGCTGAGAGCATCCCATATTTTAAAGGAAGGAATAGAAT 1560  
DB 1501 TCCTGAAGGACAGAAACAAATTTGCTGAGAGCATCCCATATTTTAAAGGAAGGAATAGAAT 1560  
QY 1561 CCGGAGATCTGGCACTGATGATGGGAGATTTTATTTCCACCTGGGGGATGCCATGAGA 1620  
DB 1561 CCGGAGATCTGGCACTGATGATGGGAGATTTTATTTTCCACCTGGGGGATGCCATGAGA 1620  
QY 1621 GGGTTGGGAACAAGAGGATTAAGTGTGATGAGCTTGGGCAAGAGAGGACACTTTG 1680  
DB 1621 GGGTTGGGAACAAGAGGATTAAGTGTGATGAGCTTGGGCAAGAGAGGACACTTTG 1680  
QY 1681 CATCTGTCTGGCAACCGCTCACTCTACAATGTGAATGAGCTGAAAGCAGACAGCTTGGTGA 1740  
DB 1681 CATCTGTCTGGCAACCGCTCACTCTACAATGTGAATGAGCTGAAAGCAGACAGCTTGGTGA 1740  
QY 1741 CCCCCAAGAAAAAGGGGCTACAGAGATTAGTAAAGCTTTAGAAAGAACTGGAAGTTAA 1800  
DB 1741 CCCCCAAGAAAAAGGGGCTACAGAGATTAGTAAAGCTTTAGAAAGAACTGGAAGTTAA 1800  
QY 1801 TCCGAGATGAGGCTTGGAGTATGATGATAAAGCCTTCTTCTCCTGCTGAGGATG 1860  
DB 1801 TCCGAGATGAGGCTTGGAGTATGATGATAAAGCCTTCTTCTCCTGCTGAGGATG 1860  
QY 1861 AAAACCTTGAGGAAAAAGGGGACTGGAGCCAGTTTCACGCTGTGGCAGCAAGGAAGAA 1920  
DB 1861 AAAACCTTGAGGAAAAAGGGGACTGGAGCCAGTTTCACGCTGTGGCAGCAAGGAAGAA 1920



Db 421 AGGCAGAACCCCAAGATATCGAAGATGACAAAGAACAAATTCAGTCCCTTCTCCATG 480  
 Qy 481 AAATGGTACACGACAGAAATGTTGAGGGAAGAACTTGCACAAAGAGATGGACCCACAG 540  
 Db 481 AAATGGTACACGACAGAAATGTTGAGGGAAGAACTTGCACAAAGAGATGGACCCACAG 540  
 Qy 541 GAGAACCAACACAGAGAGATGATGATGTTCTTATGGCGACTGATGATGATGATGATGATG 600  
 Db 541 GAGAACCAACACAGAGAGATGATGATGTTCTTATGGCGACTGATGATGATGATGATGATG 600  
 Qy 601 AGACCTTGGAACTTGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGCTGGAAGAGA 660  
 Db 601 AGACCTTGGAACTTGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGCTGGAAGAGA 660  
 Qy 661 CAGTTTCAACAGACTGTAATCAGATATGGAAGAGATGATCTGAGCAGGAAATCCAG 720  
 Db 661 CAGTTTCAACAGACTGTAATCAGGATATGGAAGAGATGATCTGAGCAGGAAATCCAG 720  
 Qy 721 ATTCAGTGAACCACTAGTAGAAGATGAAGATTTGCCACCATGATACAGATGATGATACAT 780  
 Db 721 ATTCAGTGAACCACTAGTAGAAGATGAAGATTTGCCACCATGATACAGATGATGATACAT 780  
 Qy 781 ACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAATGAAAGGATGAAATCA 840  
 Db 781 ACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAATGAAAGGATGAAATCA 840  
 Qy 841 CAGAAATGCTGCTCCCTCCTGAGGATATCTGTAAGATTTACAGGATATTTAGAG 900  
 Db 841 CAGAAATGCTGCTCCCTCCTGAGGATATCTGTAAGATTTACAGGATATTTAGAG 900  
 Qy 901 AGTAAAGCATTTTCTCTGGAAGAAACAGCAGGAGTACCAACGAAACAAATAGAAAAA 960  
 Db 901 AGTAAAGCATTTTCTCTGGAAGAAACAGCAGGAGTACCAACGAAACAAATAGAAAAA 960  
 Qy 961 CAGATGATCCAGAAACAAAGAAAGTAAAGAAAGAAAGCCTTAACCTTTAAATTAAT 1020  
 Db 961 CAGATGATCCAGAAACAAAGAAAGTAAAGAAAGAAAGCCTTAACCTTTAAATTAAT 1020  
 Qy 1021 TTGATAAGCATTTAAAGTGAACCTTGTGTCAGAAAGAACTCCGTAAGAGGGAAGAA 1080  
 Db 1021 TTGATAAGCATTTAAAGTGAACCTTGTGTCAGAAAGAACTCCGTAAGAGGGAAGAA 1080  
 Qy 1081 TTGAGGAAGCAGTGAATGATTTAAAGAACTAGTAGCAAAATACCTCAGAGTCCACAG 1140  
 Db 1081 TTGAGGAAGCAGTGAATGATTTAAAGAACTAGTAGCAAAATACCTCAGAGTCCACAG 1140  
 Qy 1141 CAAGATATGGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGTAAATGAGG 1200  
 Db 1141 CAAGATATGGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGTAAATGAGG 1200  
 Qy 1201 TGCTAGTGGAGCCATCGAGACTACCAAGAGTGGCCAGCCTACTGATGTCCCTGCAG 1260  
 Db 1201 TGCTAGTGGAGCCATCGAGACTACCAAGAGTGGCCAGCCTACTGATGTCCCTGCAG 1260  
 Qy 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAAATTTCTAGTCAATGA 1320  
 Db 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAAATTTCTAGTCAATGA 1320  
 Qy 1321 GAGGTTCCCTGCTTACCTCCGACAGATTTAGTTTCACTATTTCCCAATGATCTTCTTAA 1380  
 Db 1321 GAGGTTCCCTGCTTACCTCCGACAGATTTAGTTTCACTATTTCCCAATGATCTTCTTAA 1380  
 Qy 1381 AAAATGACCTTGGCTGGGATACCTCTTGATAGGAGATATGCAATGCAAGAAAGTTT 1440  
 Db 1381 AAAATGACCTTGGCTGGGATACCTCTTGATAGGAGATATGCAATGCAAGAAAGTTT 1440  
 Qy 1441 ATGAAGAGGTGCTGAGTGAACCTTAATGATGGCTTTGCTAAAGTCCATTTATGGCTTCA 1500  
 Db 1441 ATGAAGAGGTGCTGAGTGAACCTTAATGATGGCTTTGCTAAAGTCCATTTATGGCTTCA 1500  
 Qy 1501 TCCTGAGGCACAGAACAAATTTGCTGAGAGCATCCATATTTAAAGGAGGAATAGAT 1560  
 Db 1501 TCCTGAGGCACAGAACAAATTTGCTGAGAGCATCCATATTTAAAGGAGGAATAGAT 1560

Qy 1561 CCGGAGATCCTGGCACTGATGATGGGAGATTTATTTTCCACTGGGGGATGCCATGCAGA 1620  
 Db 1561 CCGGAGATCCTGGCACTGATGATGGGAGATTTATTTTCCACTGGGGGATGCCATGCAGA 1620  
 Qy 1621 GGGTTGGGAAACAAAGAGGCAATATAGTGGTATGAGCTTTGGGCAAGAGAGAGACACTTTG 1680  
 Db 1621 GGGTTGGGAAACAAAGAGGCAATATAGTGGTATGAGCTTTGGGCAAGAGAGAGACACTTTG 1680  
 Qy 1681 CATCTGCTGGCAACGCTCAGCTTCAATGTGAATGGACTGAAAGCAGAGCCTTGGTGA 1740  
 Db 1681 CATCTGCTGGCAACGCTCAGCTTCAATGTGAATGGACTGAAAGCAGAGCCTTGGTGA 1740  
 Qy 1741 CCCCAGAAACAAAGGCTTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
 Db 1741 CCCCAGAAACAAAGGCTTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
 Qy 1801 TCCGAGATGAAGGCTTGGAGTGATGAAGCAAGAGTCTTCTTCCCTGCTGAGGATG 1860  
 Db 1801 TCCGAGATGAAGGCTTGGAGTGATGAAGCAAGAGTCTTCTTCCCTGCTGAGGATG 1860  
 Qy 1861 AAAACCTGAGGAAAGGAGCTGAGCCAGTTTCAACGCTGTGGCAGCAAGGAAGAA 1920  
 Db 1861 AAAACCTGAGGAAAGGAGCTGAGCCAGTTTCAACGCTGTGGCAGCAAGGAAGAA 1920  
 Qy 1921 ATGAAAATGCTGCAAGGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGGAGA 1980  
 Db 1921 ATGAAAATGCTGCAAGGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGGAGA 1980  
 Qy 1981 CACAGAGTGCAGAGAGGACAGATCAAAATATTTCCATCATGCAACCCCGGAGCTCACGTGT 2040  
 Db 1981 CACAGAGTGCAGAGAGGACAGATCAAAATATTTCCATCATGCAACCCCGGAGCTCACGTGT 2040  
 Qy 2041 GGCGGCACACAGAGGCCCAAACTGCAGGCTCCGAATGCACTGGGCTTGGTGAATCCCA 2100  
 Db 2041 GGCGGCACACAGAGGCCCAAACTGCAGGCTCCGAATGCACTGGGCTTGGTGAATCCCA 2100  
 Qy 2101 AGAAGGCTCAGAGATTCGATGTGCCAAGAGACAGAGCTGGGAGGAGCAAGGTGC 2160  
 Db 2101 AGAAGGCTCAGAGATTCGATGTGCCAAGAGACAGAGCTGGGAGGAGCAAGGTGC 2160  
 Qy 2161 TCATCTTTGATGATCTCTTTGAGCAGAGGATGAGGAGGATGCTCATCTTTCCGGCTGA 2220  
 Db 2161 TCATCTTTGATGATCTCTTTGAGCAGAGGATGAGGAGGATGCTCATCTTTCCGGCTGA 2220  
 Qy 2221 TATTATCTGAGATGTGGGATCCCGAATCCGACACAGAGAGAGCGAGCTTCCAG 2280  
 Db 2221 TATTATCTGAGATGTGGGATCCCGAATCCGAGACTGACACAGAGAGAGCGAGCTTCCAG 2280  
 Qy 2281 CAATTTAGCATGATTTCAATCAAGCTTGGGAAACTCTGGAGAGA 2324  
 Db 2281 CAATTTAGCATGATTTCAATCAAGCTTGGGAAACTCTGGAGAGA 2324

RESULT 3  
 ADA00640  
 ID ADA00640 standard; cdna; 2324 BP.  
 XX  
 AC ADA00640;  
 XX  
 DT 06-NOV-2003 (first entry)  
 XX  
 DE cdna encoding human aspartyl (asparaginyl) beta-hydroxylase (HAAH).  
 XX  
 KW Tumour growth inhibition; human aspartyl (asparaginyl) beta-hydroxylase;  
 KW HAAH hydroxylation; NOTCH polypeptide;  
 KW epidermal growth factor-like repeat; EGF-like repeat; tumour cell;  
 KW malignant neoplasm; colon cancer; breast cancer; pancreatic cancer;  
 KW liver cancer; cancer of the bile duct; cancer the central nervous system;  
 KW CNS; cycostatic; gene; ss; human.  
 XX  
 OS Homo sapiens.  
 XX

PH Key Location/Qualifiers  
 FT CDS 12..2288  
 FT /\*tag= a  
 FT /product= "HAAH"  
 XX US2003031670-A1.  
 XX 13-FEB-2003.  
 XX 08-NOV-1999; 99US-00436184.  
 XX 08-NOV-1999; 99US-00436184.  
 XX (WAND/) WANDS J R.  
 XX (DMON/) DE LA MONTE S M.  
 XX (INCE/) INCE N.  
 XX (CARL/) CARLSON R I.  
 XX Wands JR, De La Monte SM, Ince N, Carlson RI;  
 XX WPI; 2003-605701/57.  
 XX P-PSDB; ADA00639.  
 XX Inhibiting tumor growth or killing tumor cells (e.g. cancer of the colon,  
 XX breast, pancreatic, liver or the central nervous system), by  
 XX administering an inhibitor of the human aspartyl (asparaginyl) beta-  
 XX hydroxylase.  
 XX Disclosure; Page 3-4; 30pp; English.  
 XX The present invention relates to a method for inhibiting tumour growth in  
 XX a mammal. The method comprises administering to the mammal a compound,  
 XX which inhibits the expression or enzymatic activity of a human aspartyl  
 XX (asparaginyl) beta-hydroxylase (HAAH). The compound may inhibit HAAH  
 XX hydroxylation of a NOTCH polypeptide. In particular, the compound may  
 XX inhibit hydroxylation of an epidermal growth factor (EGF)-like repeat  
 XX sequence in a NOTCH polypeptide. The methods are useful for inhibiting  
 XX tumour growth or killing tumour cells, or for diagnosing or  
 XX prognosticating a malignant neoplasm. In particular, the tumour or  
 XX neoplasm is colon cancer, breast cancer, pancreatic cancer, liver cancer,  
 XX cancer of the bile ducts, or cancer or tumour of the central nervous  
 XX system (CNS). The present sequence encodes HAAH.  
 XX Sequence 2324 BP; 753 A; 463 C; 625 G; 483 T; 0 U; 0 Other;  
 XX  
 XX Query Match 100.0%; Score 2324; DB 9; Length 2324;  
 XX Best Local Similarity 100.0%; Pred. No. 0;  
 XX Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 XX 1 CGGACCGTCAATGGCCCGACGCGTAAAGATGCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
 XX 1 CGGACCGTCAATGGCCCGACGCGTAAAGATGCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
 XX 61 GCTCCGCGCAGCGTACGACGAGTGGCGGCGCAGCAGCAGCGGCGGCGGCGGCGGCGG 120  
 XX 61 GCTCCGCGCAGCGTACGACGAGTGGCGGCGCAGCAGCAGCGGCGGCGGCGGCGGCGG 120  
 XX 121 AGCATGGAGACACAGAAATGGAGGAAAGCGGACCTCTCGGGAATCTCATTTCTTCAGT 180  
 XX 121 AGCATGGAGACACAGAAATGGAGGAAAGCGGACCTCTCGGGAATCTCATTTCTTCAGT 180  
 XX 181 GGTTTATGGTGAATTCATTCGCTGGGCGTCTGGACATCTGTAGCTGTGCTGTTGGTTGATC 240  
 XX 181 GGTTTATGGTGAATTCATTCGCTGGGCGTCTGGACATCTGTAGCTGTGCTGTTGGTTGATC 240  
 XX 241 TTGTTGACTATAGGAAGTTCTAGGAAATCTAGGAATCTATGATGCTGATGCTGATGCTGATG 300  
 XX 241 TTGTTGACTATAGGAAGTTCTAGGAAATCTAGGAATCTATGATGCTGATGCTGATGCTGATG 300  
 XX 301 ATTTCGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 360  
 XX 301 ATTTCGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 360

QY 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACACTGAGCCCGGAGAGCAGAGTTCCTGTGG 420  
 DB 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACACTGAGCCCGGAGAGCAGAGTTCCTGTGG 420  
 QY 421 AGGAGAACCCCGAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
 DB 421 AGGAGAACCCCGAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
 QY 481 AAATGGTACACGACAGACATGTTGAGGGGAGAGACTTGCACAAAGAAAGATGGACCCACAG 540  
 DB 481 AAATGGTACACGACAGACATGTTGAGGGGAGAGACTTGCACAAAGAAAGATGGACCCACAG 540  
 QY 541 GAGAACCAACACAGAGGATGATGAGTCTTCTTATGGCGACTGATGTAGATGATAGATTG 600  
 DB 541 GAGAACCAACACAGAGGATGATGAGTCTTCTTATGGCGACTGATGTAGATGATAGATTG 600  
 QY 601 AGACCTGGAACTGAAAGTATCTCATGAGAAACCGAGCATAGTTCACCGTGGAAAGAGA 660  
 DB 601 AGACCTGGAACTGAAAGTATCTCATGAGAAACCGAGCATAGTTCACCGTGGAAAGAGA 660  
 QY 661 CAGTTTCAAGAAGCTGTATCAGGATATGGAAGAGATGATGTCTGACGAGGAAATCCAG 720  
 DB 661 CAGTTTCAAGAAGCTGTATCAGGATATGGAAGAGATGATGTCTGACGAGGAAATCCAG 720  
 QY 721 ATTCAGTGAACCAAGTAGTAGAAGATGAAGATGCAACCATGATACAGATGATGTAACAT 780  
 DB 721 ATTCAGTGAACCAAGTAGTAGAAGATGAAGATGCAACCATGATACAGATGATGTAACAT 780  
 QY 781 ACCAAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA 840  
 DB 781 ACCAAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA 840  
 QY 841 CAGAACTACTGCTCCCTGAGGATATCTCTGTAAGATTCACAGGTAATTTGTAGAG 900  
 DB 841 CAGAACTACTGCTCCCTGAGGATATCTCTGTAAGATTCACAGGTAATTTGTAGAG 900  
 QY 901 AAGTAGCATTTTCTGTTGGAAGAACAGCAGGAGGATACCCACAGAAACAAATAGAAAA 960  
 DB 901 AAGTAGCATTTTCTGTTGGAAGAACAGCAGGAGGATACCCACAGAAACAAATAGAAAA 960  
 QY 961 CAGATGATCCAGAACAAAAAGTAAAGTAAAGAAAGAGCCTTAACTTTTAAATAAT 1020  
 DB 961 CAGATGATCCAGAACAAAAAGTAAAGTAAAGAAAGAGCCTTAACTTTTAAATAAT 1020  
 QY 1021 TTGATAAGACTATTAAGCTGAACCTGCTGCGAAGAACTCCGTAAGGAGGAAAAA 1080  
 DB 1021 TTGATAAGACTATTAAGCTGAACCTGCTGCGAAGAACTCCGTAAGGAGGAAAAA 1080  
 QY 1081 TTGAGGAAGCAGTGAATGCAATTTAAGAACTAGTAGCGAAATACCTCAGAGTCCACGAG 1140  
 DB 1081 TTGAGGAAGCAGTGAATGCAATTTAAGAACTAGTAGCGAAATACCTCAGAGTCCACGAG 1140  
 QY 1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAAATGATGAGG 1200  
 DB 1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAAATGATGAGG 1200  
 QY 1201 TGCTAGCTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCTACTGCTGCTGCTGAG 1260  
 DB 1201 TGCTAGCTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCTACTGCTGCTGCTGAG 1260  
 QY 1261 ACCCTGCGAGCTGAGTTTGAAGCGTCTGAGAGGCTCGTCAGACAGGCAAAATTTCTAGGTCATATGA 1320  
 DB 1261 ACCCTGCGAGCTGAGTTTGAAGCGTCTGAGAGGCTCGTCAGACAGGCAAAATTTCTAGGTCATATGA 1320  
 QY 1321 GAGGTTCCCTGCTTACCTCGAGAGATTAAGTTCAACTATTTCCCAATGATATCTTCTTAA 1380  
 DB 1321 GAGGTTCCCTGCTTACCTCGAGAGATTAAGTTCAACTATTTCCCAATGATATCTTCTTAA 1380  
 QY 1381 AAAATGACCTTGGGCTGGGATACCTCTTGATAGGAGATAATGACAAATGCAAGAAAGTTT 1440  
 DB 1381 AAAATGACCTTGGGCTGGGATACCTCTTGATAGGAGATAATGACAAATGCAAGAAAGTTT 1440  
 QY 1441 ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGGCTTTTGTCTAAGTCCATATGGCTTCA 1500

1441 ATGAAGAGTCTGAGTGTGACACCTAATGATGCTTTGCTAAGTCCATTATGGCTTCA 1500  
1501 TCCTGAAGGACACAGAAACAAATATGCTGAGAGCATCCCATATTTAAAGGAGGAATAGAT 1560  
1501 TCCTGAAGGACACAGAAACAAATATGCTGAGAGCATCCCATATTTAAAGGAGGAATAGAT 1560  
1561 CCGGAGATCTCTGGCACTGATGATGGGAGATTTATTTCCACCTGGGGGATGCCATCAG 1620  
1561 CCGGAGATCTCTGGCACTGATGATGGGAGATTTATTTCCACCTGGGGGATGCCATCAG 1620  
1621 GGGTTGGGAACAAAGAGGACATATAAGTGTATAGCTTGGGCAACAGAGAGGACACTTTG 1680  
1621 GGGTTGGGAACAAAGAGGACATATAAGTGTATAGCTTGGGCAACAGAGAGGACACTTTG 1680  
1681 CATCTGTCTGGCAACGCTCTACATGATGATGATGATGATGATGATGATGATGATGATG 1740  
1681 CATCTGTCTGGCAACGCTCTACATGATGATGATGATGATGATGATGATGATGATGATG 1740  
1741 CCCCAAGAGAAACGGGCTTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
1741 CCCCAAGAGAAACGGGCTTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
1801 TCCGAGATGAAGGCTTTCAGTGTATGATGAATGAAGCAAGGCTCTCTCTGCTGAGGATG 1860  
1801 TCCGAGATGAAGGCTTTCAGTGTATGATGAATGAAGCAAGGCTCTCTCTGCTGAGGATG 1860  
1861 AAAAAGGAGAAAGGAGCTGAGGAGCTTACGCTGTGACAGCAAGGAGAGAA 1920  
1861 AAAAAGGAGAAAGGAGCTGAGGAGCTTACGCTGTGACAGCAAGGAGAGAA 1920  
1921 ATGAAATGCTGCAAGAGAGCTCTTAAACCTGTACCTTACTAGAAAGTTCCCGGAGA 1980  
1921 ATGAAATGCTGCAAGAGAGCTCTTAAACCTGTACCTTACTAGAAAGTTCCCGGAGA 1980  
1981 CAACAGGATGCAAGAGAGAGACAGATCAAAATATTCATCATGACACCCCGGAGTCA 2040  
1981 CAACAGGATGCAAGAGAGAGACAGATCAAAATATTCATCATGACACCCCGGAGTCA 2040  
2041 GCGCGACACAGGCGCCCAACTGCGAGCTCCGAGTCCGAGTCTGCTGATTTCCCA 2100  
2041 GCGCGACACAGGCGCCCAACTGCGAGCTCCGAGTCCGAGTCTGCTGATTTCCCA 2100  
2101 AGGAGGCTGCAAGATTCGATGTGCAACAGAGACAGGACCTGGAGAGAGGAGTGC 2160  
2101 AGGAGGCTGCAAGATTCGATGTGCAACAGAGACAGGACCTGGAGAGAGGAGTGC 2160  
2161 TCATCTTTGATGATCTCTTTGAGCAGAGGATGAGCAGGATGCTCATCTTTCCGGCTGA 2220  
2161 TCATCTTTGATGATCTCTTTGAGCAGAGGATGAGCAGGATGCTCATCTTTCCGGCTGA 2220  
2221 TATTCATCTGATGATGCTGCGATCCGGAACCTGACACACAGAGAGAGGAGCTTCCAG 2280  
2221 TATTCATCTGATGATGCTGCGATCCGGAACCTGACACACAGAGAGAGGAGCTTCCAG 2280  
2281 CAATTTAGCATGAATTCATGCAAGTCTGGGAACTCTGGAGAGA 2324  
2281 CAATTTAGCATGAATTCATGCAAGTCTGGGAACTCTGGAGAGA 2324

RESULT 4  
ADP23914  
ID ADP23914 standard; cdna; 2452 BP.  
XX  
AC ADP23914;  
DT 18-NOV-2004 (first entry)  
XX  
KW PRO polypeptide encoding cdna SEQ ID NO:1092.  
KW ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;  
KW immunosuppressive; osteopathic; antidiabetic; dermatological;  
KW antipsoaric; antiallergic; antiasthmatic; hepatotropic; respiratory;

gene therapy; immune system.  
Unidentified.  
WO2004041170-A2.  
21-MAY-2004.  
30-OCT-2003; 2003WO-US034312.  
01-NOV-2002; 2002US-0423394P.  
(GETH ) GENENTECH INC.  
Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;  
Wu TD;  
WPI; 2004-419628/39.  
P-PSDB; ADP23915.  
New PRO polypeptides and polynucleotides, useful for treating e.g.  
erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated  
renal disease, or demyelinating diseases of the central or peripheral  
nervous system.  
Claim 1; SEQ ID NO 1092; 2940pp; English.  
The invention relates to a novel isolated nucleic acid and the PRO  
polypeptide encoded by it. A protein of the invention has  
antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,  
osteopathic, antidiabetic, dermatological, antipsoaric, antiallergic,  
antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide  
of the invention may have a use in gene therapy. The PRO polypeptide, its  
agonist, antagonist, or antibody that specifically binds to the  
polypeptide is useful for treating an immune related disorder such as  
systemic lupus erythematous, rheumatoid arthritis, osteoarthritis,  
juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis,  
idiopathic inflammatory myopathy, Sjogren's syndrome, autoimmune  
vasculitis, sarcoidosis, autoimmune haemolytic anaemia, immune-mediated renal  
thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal  
disease, a demyelinating disease of the central or peripheral nervous  
system, idiopathic demyelinating polynuropathy, Guillain-Barre syndrome,  
a chronic inflammatory demyelinating polynuropathy, a hepatobiliary  
disease, infectious or autoimmune chronic active hepatitis, primary  
biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,  
inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's  
disease, an autoimmune or immune-mediated skin disease, a bullous skin  
disease, erythema multiforme, contact dermatitis, psoriasis, an allergic  
disease, asthma, allergic rhinitis, atopic dermatitis, food  
hypersensitivity, urticaria, an immunologic disease of the lung,  
eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity  
pneumonitis, a transplantation associated disease, graft rejection or  
graft-versus-host disease. The present sequence encodes a PRO protein of  
the invention.

Sequence 2452 BP; 779 A; 507 C; 657 G; 509 T; 0 U; 0 Other;  
Query Match 99.9%; Score 2320.8; DB 13; Length 2452;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2322; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CGGACCCGTGCAATGCCAGCGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
Db 67 CGGACCCGTGCAATGCCAGCGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 126  
QY 61 GCTCGGAGCGGTAGCAGCAGTGGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAA 120  
Db 127 GCTCGGAGCGGTAGCAGCAGTGGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAA 186  
QY 121 AGCATGGAGGACACAAAGAAATGGGAGGAAGCGGACCTCTCGGGAACCTTCCTTCAGT 180  
Db 187 AGCATGGAGGACACAAAGAAATGGGAGGAAGCGGACCTCTCGGGAACCTTCCTTCAGT 246

QY 181 GGTATATGGTATGCAATTCCTGGGCGCTGGACATCTGTAGCTGTCTGTTGGTTGATC 240  
DB 247 GGTATATGGTATGCAATTCCTGGGCGCTGGACATCTGTAGCTGTCTGTTGGTTGATC 306  
QY 241 TTGTTCACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATCTCATGGTGTATGGAG 300  
DB 307 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATCTCATGGTGTATGGAG 366  
QY 301 ATTTTCACTGATGATGCCAAAGTTTTTATAGGACTTAAGAGAGACTACTTTCAGAGC 360  
DB 367 ATTTTCACTGATGATGCCAAAGTTTTTATAGGACTTAAGAGAGACTACTTTCAGAGC 426  
QY 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACTGAGCCCGAGAGCAGGTTCTGTGG 420  
DB 427 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACTGAGCCCGAGAGCAGGTTCTGTGG 486  
QY 421 AGGCAGACCCAGATATCGAAGATGAAGCAAAAGAACTTCAAGTCCCTTCTCCATG 480  
DB 487 AGGCAGACCCAGATATCGAAGATGAAGCAAAAGAACTTCAAGTCCCTTCTCCATG 546  
QY 481 AAATGTTACACGCGAACAATGTTGAGGAGAAAGACTTGCACAAAGAGATGGACCCACAG 540  
DB 547 AAATGTTACACGCGAACAATGTTGAGGAGAAAGACTTGCACAAAGAGATGGACCCACAG 606  
QY 541 GAGAACCAACAAGAGAGATGATGATGTTCTTATGGCACTGATGATGATGATGATG 600  
DB 607 GAGAACCAACAAGAGAGATGATGATGTTCTTATGGCACTGATGATGATGATGATG 666  
QY 601 AGACCTCGGAACCTGAAGTATCTCATGAAGAAACCGAGCACTAGTACCAAGTGAAGAGA 660  
DB 667 AGACCTCGGAACCTGAAGTATCTCATGAAGAAACCGAGCACTAGTACCAAGTGAAGAGA 726  
QY 661 CAGTTTCAACAAGCTGAATCAGGATATGGAAGAGATGATGTCAGCAGCAAAATCCAG 720  
DB 727 CAGTTTCAACAAGCTGAATCAGGATATGGAAGAGATGATGTCAGCAGCAAAATCCAG 786  
QY 721 ATTCAGTGAACAGTAGTAGAAGATGAAGATGTCACCACTGATGATGATGATGATG 780  
DB 787 ATTCAGTGAACAGTAGTAGAAGATGAAGATGTCACCACTGATGATGATGATGATG 846  
QY 781 ACCAAGTCTATGAGGACAGCAGTATGATGATGATGATGATGATGATGATGATGATG 840  
DB 847 ACCAAGTCTATGAGGACAGCAGTATGATGATGATGATGATGATGATGATGATGATG 906  
QY 841 CAGAGTAACTGCTCCCTGAGGATAATCTCTGTAGAGATTCACAGGTAATGTAGAG 900  
DB 907 CAGAGTAACTGCTCCCTGAGGATAATCTCTGTAGAGATTCACAGGTAATGTAGAG 966  
QY 901 AAGTAGCAATTTTCTGTGGAAGAACAGCAGGAAGTACCACAGAAACAAATAGAAAA 960  
DB 967 AAGTAGCAATTTTCTGTGGAAGAACAGCAGGAAGTACCACAGAAACAAATAGAAAA 1026  
QY 961 CAGATGATCAGAAACAAAGTTTGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
DB 1027 CAGATGATCAGAAACAAAGTTTGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1086  
QY 1021 TTGATAGACTATTAAGCTGAATGATGCTGAGAAATACTCCGTAAAAAGGGAAAAA 1080  
DB 1087 TTGATAGACTATTAAGCTGAATGATGCTGAGAAATACTCCGTAAAAAGGGAAAAA 1146  
QY 1081 TTGAGGAGCAGTGAATGATTAAGAACTAGTAGCAAAATACCTCAGAGTCCAGAG 1140  
DB 1147 TTGAGGAGCAGTGAATGATTAAGAACTAGTAGCAAAATACCTCAGAGTCCAGAG 1206  
QY 1141 CAAGATATGGAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGAGAGAGAGAG 1200  
DB 1207 CAAGATATGGAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGAGAGAGAGAG 1266  
QY 1201 TGCTAGTGAGCCATCCGAGCCTACCAAGAGTGGCGAGCCTACTGATGCTCCCTGAG 1260  
DB 1267 TGCTAGTGAGCCATCCGAGCCTACCAAGAGTGGCGAGCCTACTGATGCTCCCTGAG 1326  
QY 1261 ACTGCTGAAGCTGATTTGAAGCGTCTGCTCAGACAGGCAACAAATTTCTAGGTCATATGA 1320

DB 1327 ACTGCTGAAGCTGATTTGAGCGTCTCAGACAGGCAACAATTTCTAGGTCATATGA 1386  
QY 1321 GAGGTTCCCTGCTTACCTCGACAGATTAAGTTCAATATTTCCCAATGATCTTCTTAA 1380  
DB 1387 GAGGTTCCCTGCTTACCTCGACAGATTAAGTTCAATATTTCCCAATGATCTTCTTAA 1446  
QY 1381 AAAATGACCTTGGCGTGGGATACCTCTTGATAGAGATAATGCAATGCAAGAAAGTTT 1440  
DB 1447 AAAATGACCTTGGCGTGGGATACCTCTTGATAGAGATAATGCAATGCAAGAAAGTTT 1506  
QY 1441 ATGAAGAGGCTCTCAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTTAGGCTTCA 1500  
DB 1507 ATGAAGAGGCTCTCAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTTAGGCTTCA 1566  
QY 1501 TCCTGAAGGACAGAACAAATTTCTGAGAGCATCCCATATTTAAAGAAAGGAATAGAAT 1560  
DB 1567 TCCTGAAGGACAGAACAAATTTCTGAGAGCATCCCATATTTAAAGAAAGGAATAGAAT 1626  
QY 1561 CGGAGATTCCTGGCACTGATGATGGAGATTTTATTTTCCACTGGGGGATGCCATGSCAG 1620  
DB 1627 CGGAGATTCCTGGCACTGATGATGGAGATTTTATTTTCCACTGGGGGATGCCATGSCAG 1686  
QY 1621 GGGTTGGGAAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAACAGAGAGGACATTTG 1680  
DB 1687 GGGTTGGGAAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAACAGAGAGGACATTTG 1746  
QY 1681 CATCTGTCTGGCAACGCTCACTCTACATGTAATGCACTGAAGCAACAGCCTTGGTGA 1740  
DB 1747 CATCTGTCTGGCAACGCTCACTCTACATGTAATGCACTGAAGCAACAGCCTTGGTGA 1806  
QY 1741 CCCCAGAAAGAACCGGCTACACAGAGTTAGTAAAGTCTTTAGAAAAGAACTGGAAGTTAA 1800  
DB 1807 CCCCAGAAAGAACCGGCTACACAGAGTTAGTAAAGTCTTTAGAAAAGAACTGGAAGTTAA 1866  
QY 1801 TCCGAGATGAAGGCTTGCAGTGGATGAAGCAAAAGCTCTCTTCTCCCTGAGGATG 1860  
DB 1867 TCCGAGATGAAGGCTTGCAGTGGATGAAGCAAAAGCTCTCTTCTCCCTGAGGATG 1926  
QY 1861 AAAACCTGAGGAAAAAGGGACTGGAGCCAGTTTACGCTGTGGCAGCAAGGAGAGAA 1920  
DB 1927 AAAACCTGAGGAAAAAGGGACTGGAGCCAGTTTACGCTGTGGCAGCAAGGAGAGAA 1986  
QY 1921 ATGAAAATGCTGCAAGAGGCTCTTAAACCTGTACTTACTAGAAAAGTTTCCCGAGA 1980  
DB 1987 ATGAAAATGCTGCAAGAGGCTCTTAAACCTGTACTTACTAGAAAAGTTTCCCGAGA 2046  
QY 1981 CAAAGGATGCAAGAGAGAGAGATCAAAATATTTCCATCATGCAACCCCGGAGCTCACGTTG 2040  
DB 2047 CAAAGGATGCAAGAGAGAGAGATCAAAATATTTCCATCATGCAACCCCGGAGCTCACGTTG 2106  
QY 2041 GGCCGCAACACAGGGCCCAAACTGCAAGGCTCCGAATGCACTGGGCTTGGTGAATCCCA 2100  
DB 2107 GGCCGCAACACAGGGCCCAAACTGCAAGGCTCCGAATGCACTGGGCTTGGTGAATCCCA 2166  
QY 2101 AGGAAGGCTGCAAGATTCGATGTGCCAAAGAGACAGGACCTGGGAGAGAGAGAGTGC 2160  
DB 2167 AGGAAGGCTGCAAGATTCGATGTGCCAAAGAGACAGGACCTGGGAGAGAGAGAGTGC 2226  
QY 2161 TCATCTTTGATGATCTCTTTTGGAGCAGAGTATGCGAGGATGCTCTATTTTCCGGCTGA 2220  
DB 2227 TCATCTTTGATGATCTCTTTTGGAGCAGAGTATGCGAGGATGCTCTATTTTCCGGCTGA 2286  
QY 2221 TATTCATGCTGATGATGCTGATCCGAACTGACACCAACAGCAGAGAGAGAGAGAGTCCAG 2280  
DB 2287 TATTCATGCTGATGATGCTGATCCGAACTGACACCAACAGCAGAGAGAGAGAGAGTCCAG 2346  
QY 2281 CAATTTAGCATGAATTCATCAAGCTTTGGAAACTCTCGAGAGA 2324  
DB 2347 CAATTTAGCATGAATTCATCAAGCTTTGGAAACTCTCGAGAGA 2390



AD97347	AD97347 standard; DNA; 2452 BP.
ID	AD97347
XX	AC
XX	AD97347;
XX	02-DEC-2004 (first entry)
XX	Human ASPH DNA, an apoptosis related target Seq 55.
DE	gene; ds; human; apoptosis; cancer; inflammation; autoimmune;
KW	neurodegenerative disorder; cytostatic; antiinflammatory;
KW	immunosuppressive; neuroprotective; gene therapy; ASPH;
KW	aspartate beta hydroxylase.
XX	
OS	Homo sapiens.
XX	WO2004078783-A2.
XX	16-SEP-2004.
XX	05-MAR-2004; 2004WO-GB000957.
XX	07-MAR-2003; 2003GB-00005267.
XX	(ETRX-) EIRX THERAPEUTICS LTD.
XX	Murphy FJ, Sheehan DE, Keating KE, Hayes I, Seera L;
PI	WPI; 2004-662402/64.
XX	P-PSDB; AD97348.
DR	
DR	Identifying an agent that modulates the function of an apoptosis-
XX	associated polypeptide, useful for diagnosing or treating e.g. cancer,
PT	comprises comparing the binding of the polypeptide to the candidate agent
PT	and to a control agent.
XX	
XX	Claim 2; SEQ ID NO 55; 304pp; English.
PS	
XX	This invention relates to novel agents that modulates the function of
CC	human apoptosis-associated proteins specified within the specification.
CC	Specifically, it refers to a method for the identification of target
CC	genes whose expression is correlated with an early stage in the
CC	regulation of apoptosis. The present invention describes a method of
CC	contacting either candidate agents or control agents to the target genes
CC	and assessing the difference of binding and inhibitory activity, where
CC	the candidate agent is selected from a low molecular weight organic
CC	molecule, an antibody or fragment thereof, an antisense oligonucleotide,
CC	a small inhibitory dRNA, or a ribozyme. As such, the compositions and
CC	methods are useful for diagnosing and treating diseases or conditions
CC	associated with abnormal apoptosis in mammalian tissue, such as cancer,
CC	inflammation, autoimmune or neurodegenerative disorders. Accordingly,
CC	they exhibit cytostatic, antiinflammatory, immunosuppressive and
CC	neuroprotective activities. These may also be used for drug screening
CC	purposes and in gene therapy. This polynucleotide sequence is a human
CC	target gene associated with the regulation of apoptosis that can be
CC	modulated by novel agents of the invention.
XX	
SQ	Sequence 2452 BP; 779 A; 507 C; 657 G; 509 T; 0 U; 0 Other;
	Query Match 99.9%; Score 2320.8; DB 13; Length 2452;
	Best Local Similarity 99.9%; Pred. No. 0;
	Matches 2322; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY	1 CGAACCGTGCATGGCCACGCTAGCATGCCAAGCAGCGCGCAACAGCAGCAGCAGCG 60
Db	67 CGAACCGTGCATGGCCACGCTAGAAAGCCAAAGCAGCGCGCAACAGCAGCAGCAGCG 126
QY	61 GCTCCGGCAGCGGTAGCAGAGTGC CGGCGCAGCAGCGCCCGGGGCCCGGAGAGACAA 120
Db	127 GCTCCGGCAGCGGTAGCAGAGTGC CGGCGCAGCAGCGCCCGGGGCCCGGAGAGACAA 186
QY	121 AGCATGAGACACAGATGGAGCAAGCGGACTCTCGGAACTTCATTCTTCAGGT 180





QY	121	AGCATGAGGACACAAAGATGGAGAAAGGCGGACTCTCGGAACTTCATTCTTCCAGT	180
Db	378	AGCATGAGGACACAAAGATGGAGAAAGGCGGACTCTCAGGAACTTCATTCTTCCAGT	437
QY	181	GGTTTATGGTGATGTCCTCTGGCGCTCTGGACATCTGTAGCTGCTGTTGGTTGATC	240
Db	438	GGTTTATGGTGATGTCCTCTGGCGCTCTGGACATCTGTAGCTGCTGTTGGTTGATC	497
QY	241	TTGTTGACTATGAGGAAGTTCTAGGAAATCTAGGAATCTATGATGCTGATGGTGGAG	300
Db	498	TTGTTGACTATGAGGAAGTTCTAGGAAATCTAGGAATCTATGATGCTGATGGTGGAG	557
QY	301	ATTTTGTATGGTATGATGCGCAAGTTTATAGGACTTAAAGAGAGATCTACTTCAGGC	360
Db	558	ATTTTGTATGGTATGATGCGCAAGTTTATAGGACTTAAAGAGAGATCTACTTCAGGC	617
QY	361	CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTGTGG	420
Db	618	CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTGTGG	677
QY	421	AGGAGAACCCAGAGATATCGAAGATGAAGCAAAAGCAAAATTCAGTCCCTTCTCCATG	480
Db	678	AGGAGAACCCAGAGATATCGAAGATGAAGCAAAAGCAAAATTCAGTCCCTTCTCCATG	737
QY	481	AAATGGTACACGACAGATGTTGAGGAGAGACTTGCACAGAGAGATGGACCCACAG	540
Db	738	AAATGGTACACGACAGATGTTGAGGAGAGACTTGCACAGAGAGATGGACCCACAG	797
QY	541	GAGACCAACAAGAGATGATGATGTTCTTATGGGAGCTGATGTAGATGATAGATTG	600
Db	798	GAGACCAACAAGAGATGATGATGTTCTTATGGGAGCTGATGTAGATGATAGATTG	857
QY	601	AGACCTTGGAACTCTCATGAAAGAACCGAGCATAGTACACGCTGGAAGAGA	660
Db	858	AGACCTTGGAACTCTCATGAAAGAACCGAGCATAGTACACGCTGGAAGAGA	917
QY	720	CAGTTTTCACAGACTGATCAGGATATGGAAGAGATGATCTGAGCAGGAAATCCAG	780
Db	918	CAGTTTTCACAGACTGATCAGGATATGGAAGAGATGATCTGAGCAGGAAATCCAG	977
QY	781	ATTCAGTGAACCCAGTATGAGAGATGAAGATTCACCATGATACATGATTAACAT	1037
Db	978	ATTCAGTGAACCCAGTATGAGAGATGAAGATTCACCATGATACATGATTAACAT	1097
QY	841	CAGAGTATGAGGAAACAGCAGTATGACCTCTAGAAATGAAGGATAGAAATCA	840
Db	1038	ACCAGTCTATGAGGAAACAGCAGTATGACCTCTAGAAATGAAGGATAGAAATCA	1097
QY	841	CAGAGTATGAGGAAACAGCAGTATGACCTCTAGAAATGAAGGATAGAAATCA	900
Db	1098	CAGAGTATGAGGAAACAGCAGTATGACCTCTAGAAATGAAGGATAGAAATCA	1157
QY	901	AGTAAAGCATTTTCTCTGGAGAACAGCAGGATACCCAGAGAAACAAATAGAAAA	960
Db	1158	AGTAAAGCATTTTCTCTGGAGAACAGCAGGATACCCAGAGAAACAAATAGAAAA	1217
QY	961	CAGATGATCCAGACAAAGCAAAAGTTAAGAAAGAAAGCTTAACTTTTAAATAAT	1020
Db	1218	CAGATGATCCAGACAAAGCAAAAGTTAAGAAAGAAAGCTTAACTTTTAAATAAT	1277
QY	1021	TTGATAAGACTATTAAGAGCTGAACTTTGATGCTGCAGAAAAAATCCGCTAAAGGGGAAA	1080
Db	1278	TTGATAAGACTATTAAGAGCTGAACTTTGATGCTGCAGAAAAAATCCGCTAAAGGGGAAA	1337
QY	1081	TTGAGGAGCAGTGAATGCTTTAAAGAACTAGTACGCAAAATCCCTCAGAGTCCACGAG	1140
Db	1338	TTGAGGAGCAGTGAATGCTTTAAAGAACTAGTACGCAAAATCCCTCAGAGTCCACGAG	1397
QY	1141	CAGATATGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAAAGAGAGAGTAATGAGG	1200
Db	1398	CAGATATGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAAAGAGAGAGTAATGAGG	1457
QY	1201	TGCTACGTGGAGCATCGAGACCTTACCAAGAGTGGCCAGCCTACCTGATGCTCCCTGCAG	1260

1458	TGCTACGTGGAGCATCGAGACCTTACCAAGAGTGGCCAGCCTACCTGATGTCCTCGAG	1517
1261	ACCTGCTGAAGCTGAGTTTGAAGCGTCTCAGACAGCAAAATTTCTAGGTCAATGA	1320
1518	ACCTGCTGAAGCTGAGTTTGAAGCGTCTCAGACAGCAAAATTTCTAGGTCAATGA	1577
1321	GAGGTTCCCTCTTACCTGAGAGATTTAGTTCAACTATTTTCCCAATGATATTCCTTAA	1380
1578	GAGGTTCCCTCTTACCTGAGAGATTTAGTTCAACTATTTTCCCAATGATATTCCTTAA	1637
1381	AAAAATGACCTTGGCGTGGGATACCTCTTGTAGAGAGATTAATGACAAATGCAAGAAATTT	1440
1638	AAAAATGACCTTGGCGTGGGATACCTCTTGTAGAGAGATTAATGACAAATGCAAGAAATTT	1697
1441	ATGAAGAGTGTGATGTCGACCTTAATGATGCTTTTAAAGTCCATTTATGGCTTCA	1500
1698	ATGAAGAGTGTGATGTCGACCTTAATGATGCTTTTAAAGTCCATTTATGGCTTCA	1757
1501	TCCTCAAGGCAAGAGGATTAAGTGTGATGATGCTTGGGCAAGAGAGACATTTTG	1560
1758	TCCTCAAGGCAAGAGGATTAAGTGTGATGATGCTTGGGCAAGAGAGACATTTTG	1817
1561	CCGAGATCTCTGGCACTGATGATGAGATTTTATTCACCTGGGAGATGCAATGCA	1620
1818	CCGAGATCTCTGGCACTGATGATGAGATTTTATTCACCTGGGAGATGCAATGCA	1877
1621	GGTTTGGGAAACAAAGGATTAAGTGTGATGATGCTTGGGCAAGAGAGACATTTTG	1680
1878	GGTTTGGGAAACAAAGGATTAAGTGTGATGATGCTTGGGCAAGAGAGACATTTTG	1937
1681	CATCTGTCTGGCAACCTCACTCAATGATGATGATGATGATGATGATGATGATGATG	1740
1938	CATCTGTCTGGCAACCTCACTCAATGATGATGATGATGATGATGATGATGATGATG	1997
1741	CCCCAAAGAAAGGCTACACAGATTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA	1800
1998	CCCCAAAGAAAGGCTACACAGATTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA	2057
1801	TCCGAGATGAAGGCTTTCAGTATGATGATGATGATGATGATGATGATGATGATGATG	1860
2058	TCCGAGATGAAGGCTTTCAGTATGATGATGATGATGATGATGATGATGATGATGATG	2117
1861	AAAACTGAGGAAAGGCTGAGCAGCTTCACTGATGATGATGATGATGATGATGATG	1920
2118	AAAACTGAGGAAAGGCTGAGCAGCTTCACTGATGATGATGATGATGATGATGATG	2177
1921	ATGAAATGCTTCAAGAGGCTTCACTGATGATGATGATGATGATGATGATGATGATG	1980
2178	ATGAAATGCTTCAAGAGGCTTCACTGATGATGATGATGATGATGATGATGATGATG	2237
1981	CAACAGGATCGAAGAGGACAGATCAAAATTTTCCATCATGCAACCCGGGACTCAGGT	2040
2238	CAACAGGATCGAAGAGGACAGATCAAAATTTTCCATCATGCAACCCGGGACTCAGGT	2297
2041	GGCCGCAACAGGCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATG	2100
2298	GGCCGCAACAGGCTTCACTGATGATGATGATGATGATGATGATGATGATGATGATG	2357
2101	AGGAAGGCTGCAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATG	2160
2358	AGGAAGGCTGCAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATG	2417
2161	TCATCTTTGATGATCTTCTTTGAGCAGGATGATGATGATGATGATGATGATGATG	2220
2418	TCATCTTTGATGATCTTCTTTGAGCAGGATGATGATGATGATGATGATGATGATG	2477
2221	TATTCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2280
2478	TATTCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	2537
2281	CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTTGGAGAGA	2324

Db 2538 CAATTAGCATGAATTCATGCAAGCTTGGGAAACTCTGGAGAGA 2581

RESULT 7

ABX04178

ID ABX04178 standard; cDNA; 2680 BP.

XX

AC ABX04178;

XX

DT 10-JAN-2003 (first entry)

XX

DE Human mRNA differentially expressed in mesenchymal cells #25.

XX

KW Human; ss; gene; skeletal growth; cartilage degeneration disorder;

KW chondroblastic phenotype; mesenchymal cell; cartilage formation;

KW bone formation; arthritis; osteoarthritis; rheumatoid arthritis;

KW gout arthritis; adjuvant arthritis; arthritis deformans; anti-gout;

KW infectious arthritis; osteochondrosis; RDA; antiarthritis; osteopathic;

KW antineumatic; antiinflammatory; representational difference analysis.

XX

OS Homo sapiens.

XX

PN WO200271927-A2.

XX

PD 19-SEP-2002.

XX

PF 12-MAR-2002; 2002WO-US007787.

XX

PR 12-MAR-2001; 2001US-0274980P.

XX

PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.

XX

P1 Yates KE, Mizuno S, Glowacki J;

XX

DR WPI; 2002-723276/78.

XX

PT New nucleic acid molecules capable of promoting chondrogenesis, useful

PT for diagnosing and treating cartilaginous tissue degeneration conditions,

PT e.g. osteoarthritis, rheumatoid arthritis, gout arthritis, or

PT osteochondrosis.

XX

PS Claim 33; Page 123; 153pp; English.

XX

CC The invention relates to new isolated nucleic acid molecule comprising a

CC nucleic acid molecule consisting of a gene differentially expressed in

CC cells undergoing differentiation from mesenchymal cell to a

CC chondroblastic phenotype, or hybridising under stringent conditions to

CC them (or their fragments). Also included are expression vectors,

CC transformed host cells, expressed polypeptides or peptide fragments

CC (which induce differentiation of a mesenchymal cell and may be used as an

CC immunogen), binding partners of the polypeptides, a method for

CC identifying an agent useful in modulating mesenchymal cell

CC differentiation induction activity of a molecule, a method of diagnosing

CC a condition characterized by aberrant expression of a nucleic acid

CC molecule or its expression product; a method for determining regression,

CC progression or onset of cartilaginous tissue degeneration condition in a

CC subject characterised by aberrant expression of a nucleic acid molecule

CC or its expression product, a method for treating a cartilaginous tissue

CC degeneration condition, a method for treating a subject to reduce the

CC risk of cartilaginous tissue degeneration condition developing in the

CC subject, a method for identifying a candidate agent for treating in the

CC cartilaginous tissue degeneration condition, and a solid-phase nucleic

CC acid molecule array consisting essentially of a set of nucleic acid

CC molecule as cited above (or known from known genes shown to be

CC differentially expressed in developing mesenchymal cells using the

CC technique of representational difference analysis, RDA), its expression

CC products or fragments, fixed to a solid substrate, the nucleic acids,

CC polypeptides and agents are useful for treating cartilaginous tissue

CC degeneration conditions such as osteoarthritis, rheumatoid arthritis,

CC gout arthritis, adjuvant arthritis, arthritis deformans, infectious

CC arthritis or osteochondrosis. The present sequence is a cDNA from a

CC known gene differentially expressed in developing mesenchymal cells

XX

SQ Sequence 2680 BP; 854 A; 446 C; 575 G; 805 T; 0 U; 0 Other;

Query Match 40.7%; Score 947; DB 6; Length 2680;

Best Local Similarity 99.5%; Pred. No. 8.8e-220;

Matches 950; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGGACCGTGCATATGGCCAGCGTAAGAATGCGAAGCAGCGGCAACAGCAGCAGCAGCG 60

DB 1 CGGACCGTGCATATGGCCAGCGTAAGAATGCGAAGCAGCGGCAACAGCAGCAGCAGCG 60

QY 61 GCTCCGCGCAGCGGTAGCAGCAGTGGCGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAG 120

DB 61 GCTCCGCGCAGCGGTAGCAGCAGTGGCGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAG 120

QY 121 AGCATGAGGACACACAGAAATGGGAGGAAGGCGAAGCTCTCGGAACTTCAATCTTCACGT 180

DB 121 AGCATGAGGACACACAGAAATGGGAGGAAGGCGAAGCTCTCGGAACTTCAATCTTCACGT 180

QY 181 GGTATTAGTGATGCAATGCTGGCGCTCTGGAATCTGTAGCTGCTGTTGTTGATC 240

DB 181 GGTATTAGTGATGCAATGCTGGCGCTCTGGAATCTGTAGCTGCTGTTGTTGATC 240

QY 241 TTGTTGACTATGAGGAAGTTCTAGGAAATCTATGATCTATGATGCTGATGTTGATGAG 300

DB 241 TTGTTGACTATGAGGAAGTTCTAGGAAATCTATGATCTATGATGCTGATGTTGATGAG 300

QY 301 ATTTTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360

DB 301 ATTTTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360

QY 361 CAGCAGTCCGCGCAGGAGGCTGAGCCACACACTGAGCCCGGAGGAGAGGTTCTGTGG 420

DB 361 CAGCAGTCCGCGCAGGAGGCTGAGCCACACACTGAGCCCGGAGGAGAGGTTCTGTGG 420

QY 421 AGGCGAAGCCCGAGAAATATCGAAGATGAAGCAAAAGCAAAATTCAGTCCCTCTCTCCATG 480

DB 421 AGGCGAAGCCCGAGAAATATCGAAGATGAAGCAAAAGCAAAATTCAGTCCCTCTCTCCATG 480

QY 481 AAATGTTACACGACAGCAATGTTGAGGAGGAGCACTTTCGAAACAGCAAGATGAGCCACAG 540

DB 481 AAATGTTACACGACAGCAATGTTGAGGAGGAGCACTTTCGAAACAGCAAGATGAGCCACAG 540

QY 541 GAGAACCCACACAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600

DB 541 GAGAACCCACACAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600

QY 601 AGACCTTGAACCTGAGATATCTCATGAAGAACCGAGCATAGTTACACGTTGGAAGAGA 660

DB 601 AGACCTTGAACCTGAGATATCTCATGAAGAACCGAGCATAGTTACACGTTGGAAGAGA 660

QY 661 CAGTTTCAACAGACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720

DB 661 CAGTTTCAACAGACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720

QY 721 ATTCCAGTGAACAGTAGTAGAAGATGAAGATTCGACCATGATACAGATGATACAT 780

DB 721 ATTCCAGTGAACAGTAGTAGAAGATGAAGATTCGACCATGATACAGATGATACAT 780

QY 781 ACCAAGTCTATGAGAACAGCAGATATGAACCTCTAGAAAATGAAGGGATGAATAATCA 840

DB 781 ACCAAGTCTATGAGAACAGCAGATATGAACCTCTAGAAAATGAAGGGATGAATAATCA 840

QY 841 CAGAAAGTAACTGCTCCCTCCCTGAGGATTAATCTGTAGAAAGATTCACAGGTAAATGTAGAAG 900

DB 841 CAGAAAGTAACTGCTCCCTCCCTGAGGATTAATCTGTAGAAAGATTCACAGGTAAATGTAGAAG 900

QY 901 AAGTAAGCATTTTCTCTGTGGAAAGACAGCAGGAAAGTACCACAGAAACAAATAG 955

DB 901 AAGTAAGCATTTTCTCTGTGGAAAGACAGCAGGAAAGTACCACAGAAACAAATAG 955

RESULT 8

ADP24208



XX Labyrinthin; human; cancer; marker; antigen; detection; antibody;  
 KW MCA 44-3A6; diagnostic; vaccine; treatment; adenocarcinoma; ds.  
 XX Homo sapiens.  
 XX Key Location/Qualifiers  
 FT 70..837  
 FT /\*tag= a  
 FT /product= "lab"  
 XX WO9947683-A1.  
 XX 23-SEP-1999.  
 XX 11-MAR-1999; 99WO-US005365.  
 XX 17-MAR-1998; 98US-00040485.  
 XX (RADO/) RADOSEVICH J A.  
 XX Radosevich JA;  
 XX WPI; 1999-580307/49.  
 XX P-PSDB; AAV33642.  
 XX Novel Labyrinthin polynucleotides and polypeptides used as a diagnostic  
 PT marker for cancer and in anticancer vaccines.  
 XX Claim 1; Fig 1; 34pp; English.  
 XX This invention describes a novel polynucleotide and polypeptide sequence  
 CC for the antigen detected by the antibody MCA 44-3A6. This antigen is  
 CC designated Labyrinthin (Lab). Antibodies directed to the Labyrinthin  
 CC (Lab) protein are useful for the detection of Lab. The Lab protein is  
 CC useful in diagnostic assays for cancer, e.g. to monitor the presence and  
 CC amount of antibodies (this method is especially useful for cancer cells  
 CC that have the Lab marker). As the Lab gene is not tissue-specific, it  
 CC will detect cancer regardless of which organ it occurs in. Peptides  
 CC derived from Lab are used in the preparation of vaccines to prevent human  
 CC cancers and/or to treat humans with cancer. Antibody MCA 44-3A6 is able  
 CC to differentiate antigens associated with adenocarcinomas. However, the  
 CC sequence of the antigen detected by this antibody has not been elucidated  
 CC in the prior art. Determination of the polypeptide and polynucleotide  
 CC sequence of this antigen would enhance its usefulness in cancer  
 CC diagnosis, treatment and prevention. The present invention discloses the  
 CC sequences of the antigen recognized by the MCA 44-3A6 antibody. This  
 CC sequence encodes the human Lab protein described in the method of the  
 CC invention  
 XX  
 XX Sequence 2442 BP; 781 A; 391 C; 512 G; 758 T; 0 U; 0 Other;  
 Query Match 35.2%; Score 819; DB 2; Length 2442;  
 Best Local Similarity 98.8%; Pred. No. 1.2e-188;  
 Matches 825; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 QY 121 AGCATGAGGACACAGAAATGGAGGAAAGCGGACTCTCGGAACTTCATCTTCAGT 180  
 Db 5 AGCTTGAAGGACACAGAAATGGAGGAAAGCGGACTCTCGGAACTTCATCTTCAGT 64  
 QY 181 GGTATTATGTTGATTGCTGCGGCTCTGGAACATCTGAGCTGCTGTTGGTTGATC 240  
 Db 65 GGTATTATGTTGATTGCTGCGGCTCTGGAACATCTGAGCTGCTGTTGGTTGATC 124  
 QY 241 TTGTTGACTATGAGGAGTCTTAGGAACTTAGGAATCTATGCTGATGCTGATGAG 300  
 Db 125 TTGTTGACTATGAGGAGTCTTAGGAACTTAGGAATCTATGCTGATGCTGATGAG 184  
 QY 301 ATTTCGATGCTGATGCTGAGGAGTCTTAGGAACTTAGGAATCTATGCTGATGAG 360  
 Db 185 ATTTCGATGCTGATGCTGAGGAGTCTTAGGAACTTAGGAATCTATGCTGATGAG 244  
 QY 361 CAGCAGTCCCGCAGGAGGCTGAGCCACACACTGAGCCCGGAGGAGCAGGTTCTCTGG 420

Db 245 CAGCAGTCCCGCAGGAGGCTGAGCCACACACTGAGCCCGGAGGAGCAGGTTCTCTGG 304  
 QY 421 AGGCGAAGACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTCTCCATG 480  
 Db 305 AGGCGAAGACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTCTCCATG 364  
 QY 481 AAATGGTACACGACGAGAAACATGTTGAGCGGAGAAAGACTTCGAACAAGAGATGGACCCAG 540  
 Db 365 AAATGGTACACGACGAGAAACATGTTGAGCGGAGAAAGACTTCGAACAAGAGATGGACCCAG 424  
 QY 541 GAGAACCAACACAGAGGATGATGAGTTCCTTATCGGACATGATGATGATGATGATGATG 600  
 Db 425 GAGAACCAACACAGAGGATGATGAGTTCCTTATCGGACATGATGATGATGATGATGATG 484  
 QY 601 AGACCTTGGAACTTGAAGTATCTCATGAAGAAACCCGAGCATAGTTACCGTGGAAAGAGA 660  
 Db 485 AGACCTTGGAACTTGAAGTATCTCATGAAGAAACCCGAGCATAGTTACCGTGGAAAGAGA 544  
 QY 661 CAGTTTCAAGAGACTGTAATCAGGATATGAGGATATGAGGAGATGATGATGATGATGATG 720  
 Db 545 CAGTTTCAAGAGACTGTAATCAGGATATGAGGAGATGATGATGATGATGATGATGATG 604  
 QY 721 ATTCCAGTGAACCAAGTAGTAGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGA 780  
 Db 605 ATTCCAGTGAACCAAGTAGTAGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGA 664  
 QY 781 ACCAAGTCTATGAGGAAACGAGCATATGATGAACTCTAGAAATGAAGGAGATGAGAAATCA 840  
 Db 665 ACCAAGTCTATGAGGAAACGAGCATATGATGAACTCTAGAAATGAAGGAGATGAGAAATCA 724  
 QY 841 CAGAAGTAAGTCTGCTCCCTGAGGATATCTCTGAGGATATCTCTGAGGATATCTCTGAGG 900  
 Db 725 CAGAAGTAAGTCTGCTCCCTGAGGATATCTCTGAGGATATCTCTGAGGATATCTCTGAGG 784  
 QY 901 AAGTAAGCATTTTTCTCTGCGAAGAAACAGCAGGAAAGTACCACCCAGAAACAAATAG 955  
 Db 785 AAGTAAGCATTTTTCTCTGCGAAGAAACAGCAGGAAAGTACCACCCAGAAATCTTAAAG 839  
 RESULT 10  
 ABK39743  
 ID ABK39743 standard; cDNA; 2442 BP.  
 XX  
 AC ABK39743;  
 XX  
 DT 21-MAY-2002 (first entry)  
 XX  
 DE cDNA encoding clone #48005 (L979P) of lung tumour protein version #2.  
 XX  
 KW Lung tumour; cancer; T cell; immune response stimulator; cytostatic;  
 KW gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200204514-A2.  
 XX  
 PD 17-JAN-2002.  
 XX  
 PF 10-JUL-2001; 2001WO-US022058.  
 XX  
 PR 11-JUL-2000; 2000US-00614124.  
 PR 29-AUG-2000; 2000US-00651563.  
 PR 08-SEP-2000; 2000US-00658824.  
 PR 26-SEP-2000; 2000US-00671325.  
 PR 06-OCT-2000; 2000US-00677419.  
 PR 30-OCT-2000; 2000US-00702705.  
 PR 13-DEC-2000; 2000US-00736457.  
 PR 03-MAY-2001; 2001US-00849626.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;

PI	Marnarakis M, Carter D, Fanger GR, Vedwick TS, Bangur CS, Mcnabb A;
FI	wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;
XX	WPI; 2002-164634/21.
DR	
XX	
PT	Novel polynucleotide encoding a lung tumor polypeptide useful for
PT	stimulating and/or expanding T cells specific for a tumor protein.
XX	
PS	Claim 7; SEQ ID NO 1791; 223pp; English.
XX	
CC	The invention describes an isolated polynucleotide and polypeptide useful
CC	for stimulating and/or expanding T cells specific for a tumour protein
CC	for determining the presence of a cancer in a patient. A composition
CC	containing the polynucleotide and/or polypeptide is useful for treating a
CC	lung cancer in a patient. The polypeptide is useful for removing tumour
CC	cells from a biological sample. The polynucleotide is also useful as
CC	a probe or primer to detect the level of mRNA encoding a tumour protein.
CC	This sequence encodes a lung tumour associated protein or protein
CC	fragment, described in the method of the invention. Note: The sequence
CC	data for this patent did not form part of the printed specification, but
CC	was obtained in electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 2442 BP; 781 A; 391 C; 512 G; 758 T; 0 U; 0 Other;
	Query Match 35.2%; Score 819; DB 6; Length 2442;
	Best Local Similarity 98.8%; Pred. No. 1.2e-188;
	Matches 825; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY	121 AGCATGGAGGACACAAGAATCGGAGGAAGGCGACTCTCGGAACTTCATTCTTCA
DB	121
QY	181 GGTTTATGGTGATGCATGTCTGGGCGTCTGCA CATCTGTAGCTGTGGTTGGTTGATC
DB	181
QY	65 GGTTTATGGTGATTCATTGTCTGGGCGTCTGACATCTGTAGCTGTCTGTGGTTGATC
DB	65
QY	241 TTGTTTCACTATGAGCAAGTCTTAGCAAACCTAGGAATCTATGATGCTCATGGTGATGGAG
DB	241
QY	125 TTGTTGACTATGAGGAGTCTTAGAAAACCTAGGAATCTATGATGCTCATGGTGATGGAG
DB	125
QY	301 ATTTTGTATGGTGATGATGCCAAAGTTTTATAGGACTTAAGAGAGATCTACTTCAGAGC
DB	301
QY	185 ATTTTGTATGGTGATGATGCCAAAGTTTTATAGGACTTAAGAGAGATCTACTTCAGAGC
DB	185
QY	361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCAGAGCAGAGTTCCTGTGG
DB	361
QY	245 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCAGAGCAGAGTTCCTGTGG
DB	245
QY	421 AGGCAGAACCCAGCAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCCCATG
DB	421
QY	305 AGGCAGAACCCAGCAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCCCATG
DB	305
QY	481 AAATGTTACGCAGAACACTGTTGAGGAGAGAGACTTGCACAAAGAAGATGACCACACAG
DB	481
QY	365 AAATGGTACACGCAGAACACTGTTGAGGAGAGAGACTTGCACAAAGAAGATGACCACACAG
DB	365
QY	541 GAGAACCAACAACAGAGGATGATGAGTTCCTATGCGCATGATGATGATGATGATTTG
DB	541
QY	425 GAGAACCAACAACAGAGGATGATGAGTTCCTATGCGCATGATGATGATGATGATTTG
DB	425
QY	601 AGACCTTGGAACCTGAAGTATCTCATGNAGAAACCCAGCATCTTACCACGTGGAGAGA
DB	601
QY	485 AGACCTTGGAACCTGAAGTATCTCATGNAGAAACCCAGCATCTTACCACGTGGAGAGA
DB	485
QY	661 CAGTTTCAACAAGCTGTTAATCAGGATATGGAAGAGATGATGTTCTGAGCAGGAAAAATCCAG
DB	661
QY	721 ATTCCAGTGAACCAAGTATGAGATGAAAGATTTGCACCATGATGATACAGATGATGAACAT
DB	721
QY	605 ATTCCAGTGAACCAAGTATGAGATGAAAGATTTGCACCATGATGATACAGATGATGAACAT
DB	605



CC protein (comprising contacting T cells with the polynucleotide, protein  
 CC or antigen-presenting cells, under conditions and for a time sufficient  
 CC to permit the stimulation and/or expansion of T cells) and inhibiting the  
 CC development of a cancer in a patient (by incubating CD4<sup>+</sup> and/or CD8<sup>+</sup> T  
 CC cells isolated from a patient with the polynucleotide, protein or antigen  
 CC presenting cells that express the polynucleotide, such that T cells  
 CC proliferate, administering to the patient an effective amount of the  
 CC polynucleotide, and thus inhibiting the development of a cancer in  
 CC the patient. The polynucleotide, protein and cells are useful in a  
 CC composition for stimulating an immune response in a patient, and for  
 CC treating a cancer in a patient (particularly lung cancer). The  
 CC oligonucleotide is useful for determining the presence of a cancer in a  
 CC patient. The protein and oligonucleotides are useful in pharmaceutical  
 CC compositions, e.g. vaccines. The polynucleotide is also useful as a probe  
 CC or primer for nucleic acid hybridisation, and in the design and  
 CC preparation of ribozyme molecules for inhibiting expression of tumour  
 CC polypeptides and proteins in tumour cells. An amplified portion of the  
 CC polynucleotide is useful for isolating a full-length gene from a suitable  
 CC library. The present sequence is a cDNA (full length, extended or  
 CC partial) isolated from a library derived from lung tumour/cancer cells.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from the  
 CC USPTO at seqdata.uspto.gov/sequence.html?docid=20020197659

XX USPTO at seqdata.uspto.gov/sequence.html?docid=20020197659

XX Sequence 2442 BP; 781 A; 391 C; 512 G; 758 T; 0 U; 0 Other;  
 Query Match 35.2%; Score 819; DB 8; Length 2442;  
 Best Local Similarity 98.8%; Pred. No. 1.2e-188;  
 Matches 825; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY	121	AGCATGGAGGACACAAATGGGAGGAAAGCGGACTCTCGGAACTTCATCTTCACGT'180
DB	5	AGCTTGAAGGACACAAATGGGAGGAAAGCGGACTCTCGGAACTTCATCTTCACGT'64
QY	181	GGTTTATGTTGATTCATTCGTCGGCGTCTGGACATCTGACCTGCTGCTGTTGGTTCATC 240
DB	65	GGTTTATGTTGATTCATTCGTCGGCGTCTGGACATCTGACCTGCTGCTGTTGGTTCATC 124
QY	241	TGTTGACTATGAGGAGTTCTAGGAACTAGGAACTATGATCTGATGCTGATGTTGAG 300
DB	125	TGTTGACTATGAGGAGTTCTAGGAACTAGGAACTATGATCTGATGCTGATGTTGAG 184
QY	301	ATTTTGTATGTTGATGTCGCAAGTTTATTAGACTTAAAGAGATCTACTTCAGAGC 360
DB	185	ATTTTGTATGTTGATGTCGCAAGTTTATTAGACTTAAAGAGATCTACTTCAGAGC 244
QY	361	CAGCACTCCCGCAGAGAGGCTGAGCCACACTGAGCCCGAGGAGCAGGTTCTGTGG 420
DB	245	CAGCACTCCCGCAGAGAGGCTGAGCCACACTGAGCCCGAGGAGCAGGTTCTGTGG 304
QY	421	AGGAGAACCCCGAATATCGAAGATGAAGCAAGAAACAAATTCAGTCCCTTCCTCATG 480
DB	305	AGGAGAACCCCGAATATCGAAGATGAAGCAAGAAACAAATTCAGTCCCTTCCTCATG 364
QY	481	AAATGGTACACGAGAACATGTTGAGGAGAAAGACTTGCACCAAGAAAGATGACCCACAG 540
DB	365	AAATGGTACACGAGAACATGTTGAGGAGAAAGACTTGCACCAAGAAAGATGACCCACAG 424
QY	541	GAGAACCAACCAAGAGGATGATGATTTCTTATGCGCACTGATGTAGATGATGATTTG 600
DB	425	GAGAACCAACCAAGAGGATGATGATTTCTTATGCGCACTGATGTAGATGATGATTTG 484
QY	601	AGACCTTGGAACTGAGGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGAAGAGA 660
DB	485	AGACCTTGGAACTGAGGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGAAGAGA 544
QY	661	CAGTTTCAAGACTGTAATCAGGATATGGAAGAGATGATGTCTGAGCAGCAAAATCCAG 720
DB	545	CAGTTTCAAGACTGTAATCAGGATATGGAAGAGATGATGTCTGAGCAGCAAAATCCAG 604
QY	721	ATTCAGTGAACCACTAGTAGAAGATGAAGATTGCAACCATGATACATGATCATACAT 780
DB	605	ATTCAGTGAACCACTAGTAGAAGATGAAGATTGCAACCATGATACATGATCATACAT 664

QY	781	ACCAAGTCTATGAGGACAGGAGTATATGAACTCTAGAAAAATGAAGGATAGAAATCA 840
DB	665	ACCAAGTCTATGAGGACAGGAGTATATGAACTCTAGAAAAATGAAGGATAGAAATCA 724
QY	841	CAGAGTAACTGCTCCCTCGAGGATATCTCTGTAGAGATTCACAGTAAATTTGTAGAAG 900
DB	725	CAGAGTAACTGCTCCCTCGAGGATATCTCTGTAGAGATTCACAGTAAATTTGTAGAAG 784
QY	901	AGTAAAGCATTTTCTCTGTGGAAGAACAGCAGGAGATACCAACAGAAACAAATAG 955
DB	785	AGTAAAGCATTTTCTCTGTGGAAGAACAGCAGGAGATACCAACAGGATACCTTAAAG 839

# RESULT 12

ACA03258  
 ACA03258 standard; cDNA; 2442 BP.

XX ACA03258;  
 AC ACA03258;

XX 22-MAY-2003 (first entry)

XX Lung cancer therapyand diagnosis associated cDNA #1745.

XX Lung cancer; cytostatic; vaccine; gene therapy; cancer; gene; ss.

XX Homo sapiens.

XX US2002172952-A1.

XX 21-NOV-2002.

XX 10-JUL-2001; 2001US-00902941.

XX 30-JUN-1999; 99US-00346492.

XX 15-OCT-1999; 99US-00419356.

XX 17-DEC-1999; 99US-00466867.

XX 30-DEC-1999; 99US-00476300.

XX 06-MAR-2000; 2000US-00519642.

XX 22-MAR-2000; 2000US-00533077.

XX 10-APR-2000; 2000US-00546259.

XX 27-APR-2000; 2000US-00560406.

XX 05-JUN-2000; 2000US-00589184.

XX 11-JUL-2000; 2000US-00614124.

XX 29-AUG-2000; 2000US-00651563.

XX 08-SEP-2000; 2000US-00658824.

XX 26-SEP-2000; 2000US-00671325.

XX 06-OCT-2000; 2000US-00677419.

XX 30-OCT-2000; 2000US-00702705.

XX 13-DEC-2000; 2000US-00736457.

XX 03-MAY-2001; 2001US-00849626.

XX (CORI-) CORIXA CORP.

XX Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW;

XX Durham M, Carter D, Fanger GR, Vedvick TS, Bangur CS, McNabb A;

XX WPI; 2003-328427/31.

XX New polynucleotide, useful for preparing a composition for treating or

XX inhibiting development of cancer, e.g. lung cancer.

XX Example 5; SEQID NO 1791; 82pp; English.

XX The invention describes an isolated polynucleotide comprising one of 32  
 CC sequences, complement or degenerate variants of them. The polynucleotide  
 CC is useful for preparing a composition e.g. a vaccine or for gene therapy,  
 CC for treating or inhibiting development of cancer, e.g. lung cancer. This  
 CC sequence represents a polynucleotide associated with the compositions and  
 CC methods for the therapy and diagnosis of lung cancer

XX Sequence 2442 BP; 781 A; 391 C; 512 G; 758 T; 0 U; 0 Other;



Query Match	35.2%;	Score 819;	DB 8;	Length 2442;
Best Local Similarity	98.8%;	Pred. No. 1.2e-188;		
Matches	825;	Conservative	0;	Mismatches 10; Indels 0; Gaps 0;
QY	121	AGCATGAGGACACAGAAATGGGAGGAAAGCGGAGCTCTCGGGAACCTTCATTCTTCACGT	180	
DB	5	AGCTTGAAGGACACAGAAATGGGAGGAAAGCGGAGCTCTCAGGAATTCATTCTTCACGT	64	
QY	181	GGTTTATGCTGATGCTATGCTGGCGCTCTGGACATCTGTAGCTGCTGTTGGTTGATC	240	
DB	65	GGTTTATGCTGATGCTATGCTGGCGCTCTGGACATCTGTAGCTGCTGTTGGTTGATC	124	
QY	241	TTGTTGACTATGAGGAAGTTCTAGGAAATCTAGGAATCTATGATGCTGATGGTGTAGCGAG	300	
DB	125	TTGTTGACTATGAGGAAGTTCTAGGAAATCTATGATGCTGATGGTGTAGCGAG	184	
QY	301	ATTTTGTGATGCTGATGCTGCCAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC	360	
DB	185	ATTTTGTGATGCTGATGCTGCCAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC	244	
QY	361	CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGG	420	
DB	245	CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGG	304	
QY	421	AGGCAGAAACCCAGAGATATCGAAGATGAGCAAGAAACAAATTCAGTCCCTTCTCCATG	480	
DB	305	AGGCAGAAACCCAGAGATATCGAAGATGAGCAAGAAACAAATTCAGTCCCTTCTCCATG	364	
QY	481	AAATGCTGACACGACAGAACTGTTGAGGAGAGAACTTGCACAAAGAGATGGACCCACAG	540	
DB	365	AAATGCTGACACGACAGAACTGTTGAGGAGAGAACTTGCACAAAGAGATGGACCCACAG	424	
QY	541	GAGAACCAACAGAGGATGATGTTCTTATGCGGACTGATGATAGATGATGATGATG	600	
DB	425	GAGAACCAACAGAGGATGATGTTCTTATGCGGACTGATGATAGATGATGATGATG	484	
QY	601	AGACCTTGGAACTCAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTGGAAGAGA	660	
DB	485	AGACCTTGGAACTCAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTGGAAGAGA	544	
QY	661	CAGTTTCAAGACTGTATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG	720	
DB	545	CAGTTTCAAGACTGTATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG	604	
QY	721	ATTCAGTCAACAGTAGTAGAAGATGGAAGATTCACCATGATACAGATGATGAACAT	780	
DB	605	ATTCAGTCAACAGTAGTAGAAGATGGAAGATTCACCATGATACAGATGATGAACAT	664	
QY	781	ACCAAGTCTATGAGGAACAAAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA	840	
DB	665	ACCAAGTCTATGAGGAACAAAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA	724	
QY	841	CAGAAGTAACTGCTCCCTGAGGATAATCCTGTAGAGATTCACAGGTAATTTGAGAAG	900	
DB	725	CAGAAGTAACTGCTCCCTGAGGATAATCCTGTAGAGATTCACAGGTAATTTGAGAAG	784	
QY	901	AAGTAAAGCATTTTCTGTGGAGAGACAGCAGGAAGTACCCAGCAAAATAG	955	
DB	785	AAGTAAAGCATTTTCTGTGGAGAGACAGCAGGAAGTACCCAGCAATCTTAAAG	839	
RESULT 13				
ADH47310				
ID	ADH47310 standard; cDNA; 2442 BP.			
XX				
AC	ADH47310;			
XX				
DT	25-MAR-2004 (first entry)			
XX	Human lung tumour cDNA clone, SEQ ID No 1791.			
DE				
XX	lung cancer; tumour; cytostatic; immunostimulant; vaccine; gene therapy;			
KW	human; clone; ss.			

XX	Homo sapiens.				
OS	WO2003037267-A2.				
XX					
XX	08-MAY-2003.				
XX					
XX	28-OCT-2002; 2002WO-US034777.				
XX					
XX	29-OCT-2001; 2001US-00017754.				
PR	28-MAR-2002; 2002US-00113872.				
XX					
XX	(CORI-) CORIXA CORP.				
XX					
XX	Henderson RA, Wang T, Watanabe Y, Kalos MD, Sleath PR;				
PI	Johnson JC, Retter MW, Durham M, Carter D, Fanger GR, Vedwick TS;				
PI	Bangur CS, McNabb A;				
XX					
XX	WPI; 2003-468346/44.				
DR					
XX	New polypeptides and encoding polynucleotides, useful for diagnosing,				
PT	preventing and/or treating lung cancer.				
XX					
XX	Example 5; SEQ ID NO 1791; 258pp; English.				
PS					
XX	The invention relates to novel compositions and methods for the therapy				
CC	and diagnosis of cancer, particularly lung cancer. The compositions				
CC	comprise one or more lung tumour polypeptides, immunogenic portions				
CC	thereof, polynucleotides that encode such polypeptides, antigen presenting				
CC	cells that express such polypeptides, and T cells that are specific for				
CC	cells expressing such polypeptides. The novel compositions have				
CC	cytostatic and immunostimulant activity. The lung tumour antigens can be				
CC	used in the creation of a vaccine. The polynucleotides that encode the				
CC	lung tumour polypeptides can be used in gene therapy to help in the				
CC	treatment of lung tumors. This polynucleotide sequence represents a				
CC	human lung tumour cDNA clone of the invention. This sequence was not				
CC	shown in the specification. It has been taken from a World Intellectual				
CC	Property Organization CD ROM supplied with the specification.				
XX					
XX	Sequence 2442 BP; 781 A; 391 C; 512 G; 758 T; 0 U; 0 Other;				
Qy	Query Match	35.2%;	Score 819;	DB 10;	Length 2442;
Db	Best Local Similarity	98.8%;	Pred. NO. 1.2e-188;		
	Matches	825;	Conservative	0;	Mismatches 10; Indels 0; Gaps 0;
Qy	121	AGCATGAGGACACAGAAATGGGAGGAAAGCGGAGCTCTCGGGAACCTTCATTCTTCACGT	180		
Db	5	AGCTTGAAGGACACAGAAATGGGAGGAAAGCGGAGCTCTCAGGAATTCATTCTTCACGT	64		
Qy	181	GGTTTATGCTGATGCTATGCTGGCGCTCTGGACATCTGTAGCTGCTGTTGGTTGATC	240		
Db	65	GGTTTATGCTGATGCTATGCTGGCGCTCTGGACATCTGTAGCTGCTGTTGGTTGATC	124		
Qy	241	TTGTTGACTATGAGGAAGTTCTAGGAAATCTAGGAATCTATGATGCTGTGATGGAG	300		
Db	125	TTGTTGACTATGAGGAAGTTCTAGGAAATCTATGATGCTGTGATGGAG	184		
Qy	301	ATTTTGTGATGCTGATGCTGCAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC	360		
Db	185	ATTTTGTGATGCTGATGCTGCAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC	244		
Qy	361	CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGG	420		
Db	245	CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGG	304		
Qy	421	AGGAGAAACCCAGAGATATCGAAGATGAGCAAGAAACAAATTCAGTCCCTTCTCCATG	480		
Db	305	AGGAGAAACCCAGAGATATCGAAGATGAGCAAGAAACAAATTCAGTCCCTTCTCCATG	364		
Qy	481	AAATGTTACACGACAGAACTGTTGAGGAGAGAACTTGCACAAAGAGATGGACCCACAG	540		
Db	365	AAATGTTACACGACAGAACTGTTGAGGAGAGAACTTGCACAAAGAGATGGACCCACAG	424		



605 ATTCCAGTGAACCAAGTGTAGTGAAGATGAAGATGACCAATGATACAGATGATGTAAACAT 664

781 ACCAAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA 840

665 ACCAAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA 724

841 CAGAAGTAATGTCTCCCTGAGGATATCTCTAGAGATTCACAGGTAATGTAGAG 900

725 CAGAAGTAATGTCTCCCTGAGGATATCTCTAGAGATTCACAGGTAATGTAGAG 784

901 AAGTAAGCATTTTCTCTGGAAGAACAGCAGGAGATGACCAACAGAAAATAAG 955

785 AAGTAAGCATTTTCTCTGGAAGAACAGCAGGAGATGACCAACAGATACTTAAAG 839

RESULT 15

ACA92448

ID ACA92448 standard; DNA; 2648 BP.

XX ACA92448;

AC

XX

DT 15-JUL-2003 (first entry)

DE DNA encoding human PMM-33.

XX

XX Human; protein modification and maintenance molecule; PMM; cancer;

KW cell proliferation disorder; atherosclerosis; neurological disorder;

KW epilepsy; Huntington's disease; stroke; immune disorder; allergy;

KW inflammatory disorder; AIDS; developmental disorder; hypothyroidism;

KW Cushing's syndrome; gastrointestinal disorder; epithelial disorder;

KW infection; cytostatic; antiarteriosclerotic; anticonvulsant; nootropic;

KW neuroprotective; cerebroprotective; anti-HIV; anti-allergic; vulnary;

KW anti-inflammatory; thyromimetic; gene; ds.

XX

OS Homo sapiens.

XX

XX WO2003031939-A2.

XX

XX 17-APR-2003.

XX

XX 11-OCT-2002; 2002WO-US032850.

XX

XX 12-OCT-2001; 2001US-0329689P.

PR 25-OCT-2001; 2001US-0335703P.

PR 09-NOV-2001; 2001US-0349887P.

PR 28-NOV-2001; 2001US-0334145P.

PR 06-DEC-2001; 2001US-0337451P.

PR 14-DEC-2001; 2001US-0340584P.

XX

PA (INCY-) INCYTE GENOMICS INC.

XX

XX Ramkumar J, Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;

PI Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;

PI Sprague WW, Hafalia AJA, Chawla NK, Lehr-Watson PM, Kable AE, Yue H;

PI Marquis JP, Yao MG, Richardson TW, Tang TY, Jin P, Chien D;

PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;

XX

XX WPI; 2003-430274/40.

DR P-PSDB; AB092053.

XX

XX New human protein modification and maintenance molecules (PMM), useful

PT for diagnosing, treating and preventing diseases or conditions associated

PT with the aberrant PMM expression e.g. cancer, atherosclerosis, or

PT infections.

XX

XX Claim 5; Page 306; 311pp; English.

PS

XX The present invention relates to the isolation of human protein

CC modification and maintenance molecules (PMM), and the polynucleotide

CC sequences encoding them. A total of 40 PMM polypeptides (designated PMM

CC -1 to PMM-40) are disclosed. The sequences of the invention are useful

CC for diagnosing a condition or disease associated with the expression of

CC PMM in a subject, preparing a polyclonal or monoclonal antibody, and

generating an expression profile of a sample containing the

polynucleotides. The diseases or conditions associated with decreased

expression or overexpression of PMM are cell proliferation disorders

(e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,

Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,

allergies), developmental disorders (e.g. hypothyroidism, Cushing's

syndrome), gastrointestinal or epithelial disorders, and infections. The

PMM polypeptides or their fragments are useful in screening compounds

for effectiveness as agonists or antagonists of the polypeptides, or in

altering the expression of the target polynucleotide and compounds that

specifically bind to, or modulate the activity of the polypeptide.

ACA92416-ACA92455 encode the human PMM polypeptides of the invention

XX

SQ Sequence 2648 BP; 844 A; 420 C; 545 G; 839 T; 0 U; 0 Other;

Query Match 30.6%; Score 711.4; DB 10; Length 2648;

Best Local Similarity 92.5%; Pred. No. 2e-162;

Matches 782; Conservative 0; Mismatches 6; Indels 57; Gaps 1;

QY 111 AGAGAGACAAAGCATGGAGGACACAAGAAATGGAGGAAGCGGACTTCGGAACTTCA 170

DB 128 AAAGAGACAAAGCATGGAGGACACAAGAAATGGAGGAAGCGGACTTCGGAACTTCA 187

QY 171 TTCTTCACTGCTTTATGCTGATTCATTTGCTGGCGCTCTGGACATCTGTAGCTCGTT 230

DB 188 TTCTTCACTGCTTTATGCTGATTCATTTGCTGGCGCTCTGGACATCTGTAGCTCGTT 247

QY 231 TGGTTTGAATCTTTGCTGATCTATGAGGAAGTTCTAGGAAAATCTAGGAATCTATGATCTGAT 290

DB 248 TGGTTTGAATCTTTGCTGATCTATGAGGAAGTTCTAGGAAAATCTATGATCTGAT 307

QY 291 GGTGATGGAGATTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 350

DB 308 GGTGATGGAGATTTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 367

QY 351 ACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAG 410

DB 368 ACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAG 427

QY 411 GTTCTCTGTGAGGAGCAAAACCCAGAAATATCGAAGATGAAGCAAAAGAAATTCAGTCC 470

DB 428 GTTCTCTGTGAGGAGCAAAACCCAGAAATATCGAAGATGAAGCAAAAGAAATTCAGTCC 487

QY 471 CTTTCTCCATGAATGTGTACAGCAGAACATGTTGAGGAGAGAGACTTGCACAAAGAGAT 530

DB 488 CTTTCTCCATGAATGTGTACAGCAGAACATGTTGAGGAGAGAGACTTGCACAAAGAGAT 547

QY 531 GGACCCACAGAGAACCAACCAAGAGGATGATGATGATGATGATGATGATGATGATGAT 590

DB 548 GGACCCACAGAGAACCAACCAAGAGGATGATGATGATGATGATGATGATGATGATGAT 607

QY 591 GATAGATTGAGACCTTGGAACTTGAAGTATCTCATGAGAAACCCAGCATAGTTACCAC 650

DB 608 GATAGATTGAGACCTTGGAACTTGAAGTATCTCATGAGAAACCCAGCATAGTTACCAC 667

QY 651 GTGGAAGAGAGATTTTTCACAAAGACTGTAAATCAGGATATGGAAGAGATGATGTCGAGCAG 710

DB 668 GTGGAAGAGATTTTTCACAAAGACTGTAAATCAGGATATGGAAGAGATGATGTCGAGCAG 677

QY 711 GAAATATCCAGATTCAGTGAAACAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG 770

DB 678 -----CAGATTCAGTGAAACAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG 730

QY 771 GATGTAAATACCAAGTCTATGAGGAACAGCAGTATATGAACTCTAGAAAATGAAGGG 830

DB 731 GATGTAAATACCAAGTCTATGAGGAACAGCAGTATATGAACTCTAGAAAATGAAGGG 790

QY 831 ATAGAAATACCAAGTCTATGAGGAACAGCAGTATATGAACTCTAGAAAATGAAGGGTA 890

DB 791 ATAGAAATACCAAGTCTATGAGGAACAGCAGTATATGAACTCTAGAAAATGAAGGGTA 850

QY 891 ATTGTAGAAAGTAAAGATTTTCTGTTGGAAGAACAGCAGGAGGAGTACCAACCAAGAAACA 950

Mon Mar 28 06:06:59 2005

us-09-436-184-3.rng

Page 19

Db 851 ATTGTAGAAAGTAAGCATTTTCTGTGCGAAGAACAGCAGGAAGTACCACCAGATACT 910  
Qy 951 AATAG 955  
Db 911 TAAAG 915

Search completed: March 25, 2005, 09:48:25  
Job time : 1164 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 12:32:18 ; Search time 1232 Seconds  
(without alignments)  
11241.467 Million cell updates/sec

Title: US-09-436-184-3  
Perfect score: 2324  
Sequence: 1 CGGACGTCGAATGCGCCAG.....CTGGGAACTCTGGAGAGA 2324

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5552208 seqs, 2979665951 residues

Total number of hits satisfying Chosen parameters: 11104416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA:\*\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/1/pubpna/US06\_PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:  
19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:  
20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:  
21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:  
22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2324	100.0	2324	9	US-09-903-248-3
2	2324	100.0	2324	9	US-09-859-604-3
3	2324	100.0	2324	9	US-09-903-063-3
4	2324	100.0	2324	9	US-09-903-216-3
5	2324	100.0	2324	9	US-09-903-199-3
6	2324	100.0	2324	9	US-09-903-023-3
7	2324	100.0	2324	10	US-09-436-184-3
8	2319.2	99.8	5258	15	US-10-084-817-63
9	947	40.7	2680	16	US-10-096-534-37
10	947	40.7	2680	16	US-10-289-909-3
11	819	35.2	2442	9	US-09-736-457-1791

12	819	35.2	2442	9	US-09-902-941-1791	Sequence 1791, Ap
13	819	35.2	2442	9	US-09-849-626-1791	Sequence 1791, Ap
14	819	35.2	2442	14	US-10-017-754-1791	Sequence 1791, Ap
15	819	35.2	2442	16	US-10-113-872-1791	Sequence 1791, Ap
16	819	35.2	2442	17	US-10-283-017-1791	Sequence 1791, Ap
17	617.6	26.6	3110	14	US-10-198-846-9938	Sequence 9938, Ap
18	464.6	20.0	502	10	US-09-918-995-34279	Sequence 34279, A
19	416.2	17.9	660	11	US-09-969-034-3202	Sequence 3202, A
20	364.4	15.7	366	17	US-10-242-535A-25661	Sequence 25661, A
21	364.4	15.7	366	17	US-10-085-783A-25661	Sequence 25661, A
22	317	13.6	495	14	US-10-066-543-277	Sequence 277, App
23	170.6	7.3	1382	17	US-10-443-622-24	Sequence 24, Appl
24	153.2	6.6	161	16	US-10-029-386-26964	Sequence 26964, A
25	153.2	6.6	592	16	US-10-039-386-13264	Sequence 13264, A
26	135	5.8	151	17	US-10-242-535A-2707	Sequence 2707, Ap
27	135	5.8	151	17	US-10-085-783A-2707	Sequence 2707, Ap
28	127.4	5.5	583	11	US-09-969-034-3266	Sequence 3266, Ap
29	96.8	4.2	671	10	US-09-991-936-1528	Sequence 1528, Ap
30	84.2	3.6	554	13	US-10-101-487-69	Sequence 69, Appl
31	83.4	3.6	554	13	US-10-101-487-106	Sequence 106, Appl
32	82.8	3.6	522	13	US-10-101-487-71	Sequence 71, Appl
33	82.8	3.6	530	13	US-10-101-487-73	Sequence 73, Appl
34	78.4	3.4	720	13	US-10-101-487-76	Sequence 76, Appl
35	78.2	3.4	720	13	US-10-101-487-74	Sequence 74, Appl
36	78	3.4	450	13	US-10-027-632-2825	Sequence 2825, Ap
37	78	3.4	450	17	US-10-027-632-2825	Sequence 2825, Ap
38	74.6	3.2	17137	15	US-10-311-455-163	Sequence 163, App
39	72.2	3.1	31124	13	US-10-087-192-463	Sequence 463, App
40	71.2	3.1	3912	17	US-10-282-122A-9928	Sequence 9928, App
41	66	2.8	453	13	US-10-027-632-56205	Sequence 56205, A
42	66	2.8	453	13	US-10-027-632-310460	Sequence 310460, A
43	66	2.8	453	17	US-10-027-632-56205	Sequence 56205, A
44	66	2.8	453	17	US-10-027-632-310460	Sequence 310460, A
45	66	2.8	68233	17	US-10-034-650-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1  
US-09-903-248-3  
; Sequence 3, Application US/0903248  
; Patent No. US20020102263A1  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Roif I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV5  
; CURRENT APPLICATION NUMBER: US/09/903,248  
; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-903-248-3

Query Match	100.0%	Score 2324;	DB 9;	Length 2324;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2324;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	CGGACCGTCGAATGCGCCAGCGTAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCGG	60	
Db	1	CGGACCGTCGAATGCGCCAGCGTAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCGG	60	
QY	61	GCTCCGCGCAGCGTACAGAGTGGCGGCGAGCAGCGCCCGGGCCCGGAGAGACAA	120	
Db	61	GCTCCGCGCAGCGTACAGAGTGGCGGCGAGCAGCGCCCGGGCCCGGAGAGACAA	120	

121 AGCATGGAGGACACAAGAAATGGGAGGAAGCGGACTCTCGGGAACCTTCACTTCCAGT 180  
121 AGCATGGAGGACACAAGAAATGGGAGGAAGCGGACTCTCGGGAACCTTCACTTCCAGT 180  
181 GGTATATGGTGAATGCAATGCTGGGCGTCTGGACATCTGTAGCTGTGTTGGTTGATC 240  
181 GGTATATGGTGAATGCAATGCTGGGCGTCTGGACATCTGTAGCTGTGTTGGTTGATC 240  
241 TTGTTGACTATGAGGAAGTTCTAGGAATACTAGAAATCTATGATGCTGATGCTGATGGAG 300  
241 TTGTTGACTATGAGGAAGTTCTAGGAATACTAGGAATACTATGATGCTGATGCTGATGGAG 300  
301 ATTTTGTATGTTGATGATGCCAAAGTTTATTTAGGACTTTAAAGAGAGATCTACTTTCAGAGC 360  
301 ATTTTGTATGTTGATGATGCCAAAGTTTATTTAGGACTTTAAAGAGAGATCTACTTTCAGAGC 360  
361 CAGCAGTCCGCCAGAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGG 420  
361 CAGCAGTCCGCCAGAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGG 420  
421 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG 480  
421 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG 480  
481 AAATGGTACACGACAGAACTGTTGAGGAGAGAACTTGCACCAAGAAAGATGGACCCACAG 540  
481 AAATGGTACACGACAGAACTGTTGAGGAGAGAACTTGCACCAAGAAAGATGGACCCACAG 540  
541 GAGAACCAACAAGAGAGATGATGAGTTCTTATGCGGAGCTGATGATGATGATGATG 600  
541 GAGAACCAACAAGAGAGATGATGAGTTCTTATGCGGAGCTGATGATGATGATGATG 600  
601 AGACCTCGGAACCTGAAGTATCTCATGAAGAACCGAGCATAGTTACACGTGGGAAGAGA 660  
601 AGACCTCGGAACCTGAAGTATCTCATGAAGAACCGAGCATAGTTACACGTGGGAAGAGA 660  
661 CAGTTTCCACAGACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720  
661 CAGTTTCCACAGACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720  
721 ATTCCAGTGAACCAAGTAGTAGAAGATGAAGATTTGACCATGATACAGATGATGATACAT 780  
721 ATTCCAGTGAACCAAGTAGTAGAAGATGAAGATTTGACCATGATACAGATGATGATACAT 780  
781 ACCAAGTCTATGAGGAACAAGCAGTATATGAACTCTAGAAAATGAAGGATAGAAATCA 840  
781 ACCAAGTCTATGAGGAACAAGCAGTATATGAACTCTAGAAAATGAAGGATAGAAATCA 840  
841 CAGAAGTAACCTGCTCCCTGAGGATAATCCTGTAGAAGATTTCACAGGTAATTTGAGAG 900  
841 CAGAAGTAACCTGCTCCCTGAGGATAATCCTGTAGAAGATTTCACAGGTAATTTGAGAG 900  
901 AAGTAAGCATTTTCTGTGGAAGAACAGCAGGAAGTACCAACCAAGAAACAAATAGAAAAA 960  
901 AAGTAAGCATTTTCTGTGGAAGAACAGCAGGAAGTACCAACCAAGAAACAAATAGAAAAA 960  
961 CAGATGATCCAGAACAAAAAGTTAAGAAAAAGAGCCCTTAAATTAAT 1020  
961 CAGATGATCCAGAACAAAAAGTTAAGAAAAAGAGCCCTTAAATTAAT 1020  
1021 TTGATAAGACTATTAAAGCTGAACTTGATGCTGCAGAAAAAATCCGTAAAAAGGGAAAAA 1080  
1021 TTGATAAGACTATTAAAGCTGAACTTGATGCTGCAGAAAAAATCCGTAAAAAGGGAAAAA 1080  
1081 TTGAGGAGCAGTGAATGATTTAAAGACTAGTACCGAATACCTTCAGAGTCCAGAG 1140  
1081 TTGAGGAGCAGTGAATGATTTAAAGACTAGTACCGAATACCTTCAGAGTCCAGAG 1140  
1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAGTAATGAGG 1200  
1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAGTAATGAGG 1200

1201 TGCTACGTGGAGGCATCGAGACCTACCAAGAGGTGGCCAGCCTACCTGATGTCTCTGCAG 1260  
1201 TGCTACGTGGAGGCATCGAGACCTACCAAGAGGTGGCCAGCCTACCTGATGTCTCTGCAG 1260  
1261 ACCTGCTGAGCTGAGCTTTGAAGCGTCTGACACAGGCAACAATTTCTAGGTCATATGA 1320  
1261 ACCTGCTGAGCTGAGCTTTGAAGCGTCTGACACAGGCAACAATTTCTAGGTCATATGA 1320  
1321 GAGGTTCCCTGCTTACCTTCGACAGATTAGTTCAACTATTTCCCAATGATACTTCTCTAA 1380  
1321 GAGGTTCCCTGCTTACCTTCGACAGATTAGTTCAACTATTTCCCAATGATACTTCTCTAA 1380  
1381 AAAATGACCTTGGCGTGGGATACCTCTGTAGTAGAGATATGACATATGCAAGAAAGTTT 1440  
1381 AAAATGACCTTGGCGTGGGATACCTCTGTAGTAGAGATATGACATATGCAAGAAAGTTT 1440  
1441 ATCAAGAGGTCTGAGTGTGACACCTTAATGATGCTTAAAGTCCATTTATGGCTTCA 1500  
1441 ATCAAGAGGTCTGAGTGTGACACCTTAATGATGCTTAAAGTCCATTTATGGCTTCA 1500  
1501 TCCTGAAGGCACAGAAACAAAATTCCTGAGAGCATCCCATATTTAAAGGAAAGGAATAGAAT 1560  
1501 TCCTGAAGGCACAGAAACAAAATTCCTGAGAGCATCCCATATTTAAAGGAAAGGAATAGAAT 1560  
1561 CCGGAGATCTGCGCATGTATGATGGGAGATTTATTTCCACCTGGGGGATGCCATGCAGA 1620  
1561 CCGGAGATCTGCGCATGTATGATGGGAGATTTATTTCCACCTGGGGGATGCCATGCAGA 1620  
1621 GGGTGGGAAACAAGAGGCATATAAGTGTGATGAGCTTGGGCAACAAGAGAGGACACTTTG 1680  
1621 GGGTGGGAAACAAGAGGCATATAAGTGTGATGAGCTTGGGCAACAAGAGAGGACACTTTG 1680  
1681 CATCTGTCTGGCAACGCTCACTCTCAATGTGAATGAGCTGAAAGCAACAGCCTTGGTGA 1740  
1681 CATCTGTCTGGCAACGCTCACTCTCAATGTGAATGAGCTGAAAGCAACAGCCTTGGTGA 1740  
1741 CCCAAAAGAAACCGGCTACACAGATTTAGTAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
1741 CCCAAAAGAAACCGGCTACACAGATTTAGTAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
1801 TCCGAGATGAAGCCCTTGCAGTGTGATGATAAGCCAAAGGCTCTTCTCCCTGAGGATG 1860  
1801 TCCGAGATGAAGCCCTTGCAGTGTGATGATAAGCCAAAGGCTCTTCTCCCTGAGGATG 1860  
1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTCACGCTGTGGCAGCAAGGAAGAA 1920  
1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTCACGCTGTGGCAGCAAGGAAGAA 1920  
1921 ATGAAAATGCTGCAAGAGGAGCTCTTAAACCTGTACTTACTAGAAAAGTTCCCCGAGA 1980  
1921 ATGAAAATGCTGCAAGAGGAGCTCTTAAACCTGTACTTACTAGAAAAGTTCCCCGAGA 1980  
1981 CAACAGAGTSCAGAAGAGGACAGATCAAAATATTTCCATCATGCAACCCCGGACTCACGTGT 2040  
1981 CAACAGAGTSCAGAAGAGGACAGATCAAAATATTTCCATCATGCAACCCCGGACTCACGTGT 2040  
2041 GGCCGCAACACAGGCCCCACAACTGCAAGGCTCCGAATGCAACCTGGGCTTGGTGAATCCCA 2100  
2041 GGCCGCAACACAGGCCCCACAACTGCAAGGCTCCGAATGCAACCTGGGCTTGGTGAATCCCA 2100  
2101 AGGAAGCTCAAGATTCGATGTGCCAACGAGACCCAGGACCTGGGAGGAAGGCAAGGTGC 2160  
2101 AGGAAGCTCAAGATTCGATGTGCCAACGAGACCCAGGACCTGGGAGGAAGGCAAGGTGC 2160  
2161 TCATCTTTGATGATCTCTTTGAGCAGAGGATGAGGAGGATGCTCATCTTTCCGGCTGA 2220  
2161 TCATCTTTGATGATCTCTTTGAGCAGAGGATGAGGAGGATGCTCATCTTTCCGGCTGA 2220  
2221 TATTTCATGTTGGATGTGGCATCCCGAATCTGACACCAAGCAGAGACGAGCCTTCCAG 2280  
2221 TATTTCATGTTGGATGTGGCATCCCGAATCTGACACCAAGCAGAGACGAGCCTTCCAG 2280  
2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTCTGGAGAGA 2324





Qy	1741	CCCCAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAAACTCGAAGTTAA	1800
Db	1741	CCCCAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAAACTCGAAGTTAA	1800
Qy	1801	TCCGAGATGAAGGCCTTGCGAGTGAATGAATAAGCCAAAGTCTCTTCTGCTGAGGATG	1860
Db	1801	TCCGAGATGAAGGCCTTGCGAGTGAATGAATAAGCCAAAGTCTCTTCTGCTGAGGATG	1860
Qy	1861	AAAACTGAGGGGAAAAAGGGGACTGGAGCCAGTTACGCTGTGGCAGCAGGAAGGAAGAA	1920
Db	1861	AAAACTGAGGGGAAAAAGGGGACTGGAGCCAGTTACGCTGTGGCAGCAGGAAGGAAGAA	1920
Qy	1921	ATGAANAATGCTGCAAGAGGACTCCTAAAACCTGTACCTTACTAGAAAAGTCTCCCGAGA	1980
Db	1921	ATGAANAATGCTGCAAGAGGACTCCTAAAACCTGTACCTTACTAGAAAAGTCTCCCGAGA	1980
Qy	1981	CAACAGGATGCAGAAGAGGACAGATCAAAATATTCATCATGCACCCCGGAGCTCACGTGT	2040
Db	1981	CAACAGGATGCAGAAGAGGACAGATCAAAATATTCATCATGCACCCCGGAGCTCACGTGT	2040
Qy	2041	GGCCGCAACAGGGCCCCAAACTGCAGGCTCCGAATGCACTTGGGCTTGGTGATTTCCCA	2100
Db	2041	GGCCGCAACAGGGCCCCAAACTGCAGGCTCCGAATGCACTTGGGCTTGGTGATTTCCCA	2100
Qy	2101	AGGAAGCTGCAAGATTGATGTGCCAAACAGACCCAGGACCTGGGAGGAGGCAAGGTGC	2160
Db	2101	AGGAAGCTGCAAGATTGATGTGCCAAACAGACCCAGGACCTGGGAGGAGGCAAGGTGC	2160
Qy	2161	TCATCTTTTGATGACTCTTTTGAGCACAGAGTATGGCAGGATGCCTCATCTTTCCGGCTGA	2220
Db	2161	TCATCTTTTGATGACTCTTTTGAGCACAGGATATGGCAGGATGCCTCATCTTTCCGGCTGA	2220
Qy	2221	TATTTATCTGTGGATGTGTGGCATCCGGAACTGACCAACAGCAGAGACGAGCCTTCCAG	2280
Db	2221	TATTTATCTGTGGATGTGTGGCATCCGGAACTGACCAACAGCAGAGACGAGCCTTCCAG	2280
Qy	2281	CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA	2324
Db	2281	CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA	2324

Qy	61	GCTCCGGCAGCGGTAGCAGAGTCCGGCAGCAGCAGCCCGGGCCCGGAGAGACAA	120
Db	61	GCTCCGGCAGCGGTAGCAGAGTCCGGCAGCAGCAGCCCGGGCCCGGAGAGACAA	120
Qy	121	AGCATGGAGGACACAAGAAATGGAGGAAAGGCGGACTCTCGGGAACCTTCAATCTTCACGT	180
Db	121	AGCATGGAGGACACAAGAAATGGAGGAAAGGCGGACTCTCGGGAACCTTCAATCTTCACGT	180
Qy	181	GGTTTATGGTGATGCATTTGCTGGCGCTGTGACATCTGTAGCTGTCTGTTGGTTTGATC	240
Db	181	GGTTTATGGTGATGCATTTGCTGGCGCTGTGACATCTGTAGCTGTCTGTTGGTTTGATC	240
Qy	241	TTGTTTGACTATCAGGAAGTTCTAGGAAACTAGGAATCTATCATCTGATCTGATGCTGATGGAG	300
Db	241	TTGTTTGACTATCAGGAAGTTCTAGGAAACTAGGAATCTATCATCTGATCTGATGCTGATGGAG	300
Qy	301	ATTTTGTGATGGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC	360
Db	301	ATTTTGTGATGGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC	360
Qy	361	CAGCAGTCCGCGCAGAAAGGCTGAGCCACACACTGAGCCCGAGGAGGAGGTTCTCTGTG	420
Db	361	CAGCAGTCCGCGCAGAAAGGCTGAGCCACACACTGAGCCCGAGGAGGAGGTTCTCTGTG	420
Qy	421	AGGCAGAAACCCAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG	480
Db	421	AGGCAGAAACCCAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG	480
Qy	481	AAATGGTACACGCAGAAACATTTGTTAGGGAGAGAGACTTGCAACAAAGAAAGTGAAGCCACAG	540
Db	481	AAATGGTACACGCAGAAACATTTGTTAGGGAGAGAGACTTGCAACAAAGAAAGTGAAGCCACAG	540
Qy	541	GAGAAACCAACAAAGAGGATGATGAGTTTCTTATGGCGACTGATGTAGATGATAGATTG	600
Db	541	GAGAAACCAACAAAGAGGATGATGAGTTTCTTATGGCGACTGATGTAGATGATAGATTG	600
Qy	601	AGACCTTGGAAACCTGAAGTATCTCATGAAAGAAACCGAGCATAGTTTACCACTGGGAAGAGA	660
Db	601	AGACCTTGGAAACCTGAAGTATCTCATGAAAGAAACCGAGCATAGTTTACCACTGGGAAGAGA	660
Qy	661	CAGTTTTCACAAAGCTCTTAATCAGGATATCGAAGAGATGATGCTGAGCAGGAAAAATCCAG	720
Db	661	CAGTTTTCACAAAGCTCTTAATCAGGATATCGAAGAGATGATGCTGAGCAGGAAAAATCCAG	720
Qy	721	ATTCCAGTGAACCAAGTAGTAGAAGATGAAGATTTGCAACCATGATACAGATGATGTAAAT	780
Db	721	ATTCCAGTGAACCAAGTAGTAGAAGATGAAGATTTGCAACCATGATACAGATGATGTAAAT	780
Qy	781	ACCAAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA	840
Db	781	ACCAAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA	840
Qy	841	CAGAAGTAACTGCTCCCTCAGGATAATCCTGTAGAAGATTCACAGGTAATTTGTAGAAG	900
Db	841	CAGAAGTAACTGCTCCCTCAGGATAATCCTGTAGAAGATTCACAGGTAATTTGTAGAAG	900
Qy	901	AAGTAAAGCATTTTCTGTGTGGAAGAACAGCAGAGAGATGACCAACAGAAACAAATAGAAAA	960
Db	901	AAGTAAAGCATTTTCTGTGTGGAAGAACAGCAGAGAGATGACCAACAGAAACAAATAGAAAA	960
Qy	961	CAGATGATCCAGAACAAAAAGTAAAGAAAAAGAACCTTAAACTTTTAAATAAAT	1020
Db	961	CAGATGATCCAGAACAAAAAGTAAAGAAAAAGAACCTTAAACTTTTAAATAAAT	1020
Qy	1021	TTGATAAGACTATTTAAAGCTGAACCTTGATCTGCAGAAAAAATCCTCGTAAAGGGGAAAA	1080
Db	1021	TTGATAAGACTATTTAAAGCTGAACCTTGATCTGCAGAAAAAATCCTCGTAAAGGGGAAAA	1080
Qy	1081	TTGAGGAAGCAGTGAATGCAATTTAAAGAACTTAGTACGCAAAATACCTTCAGAGTCCACGAG	1140
Db	1081	TTGAGGAAGCAGTGAATGCAATTTAAAGAACTTAGTACGCAAAATACCTTCAGAGTCCACGAG	1140
Qy	1141	CAAGATATGGGAAGGCGCAGTGTGAGGATGATTTGGCTGTAGAGAGAGAGAAATATGAGG	1200

Db 1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTGGCTGAGAAGGAGGAAGTAATGAGG 1200  
Qy 1201 TGCTAGCTGGAGCCCATCAGACCTACCAAGAGGTGGCAGGCTACCTGATGTCCTCGCAG 1260  
Db 1201 TGCTAGCTGGAGCCCATCAGACCTACCAAGAGGTGGCAGGCTACCTGATGTCCTCGCAG 1260  
Qy 1261 ACTGCTGAAGCTGAGTTTGAAGCGTGGCTCAGACAGGCAACATTTCTAGGTCATATGA 1320  
Db 1261 ACTGCTGAAGCTGAGTTTGAAGCGTGGCTCAGACAGGCAACATTTCTAGGTCATATGA 1320  
Qy 1321 GAGGTTCCCTGTACCTCGCAGAGATTAGTTCAACTATTTCCCAATGATACTTCCCTTAA 1380  
Db 1321 GAGGTTCCCTGTACCTCGCAGAGATTAGTTCAACTATTTCCCAATGATACTTCCCTTAA 1380  
Qy 1381 AAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATAATGACAATGCCAAGAAAGTTT 1440  
Db 1381 AAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATAATGACAATGCCAAGAAAGTTT 1440  
Qy 1441 ATGAAGAGTGTGAGTGTGACACCTAAATGATGGCTTTGCTAAAGTCCATATCGCTTCA 1500  
Db 1441 ATGAAGAGTGTGAGTGTGACACCTAAATGATGGCTTTGCTAAAGTCCATATCGCTTCA 1500  
Qy 1501 TCCTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCCATATTTAAAGGAAGGAATAGAT 1560  
Db 1501 TCCTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCCATATTTAAAGGAAGGAATAGAT 1560  
Qy 1561 CCGGAGATCCTGGCAGCTGATGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGA 1620  
Db 1561 CCGGAGATCCTGGCAGCTGATGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGA 1620  
Qy 1621 GAGTTGGGAAACAAAGAGGCATATAAGTGTATGAGCTTGGGCAACAGAGAGCACTTTG 1680  
Db 1621 GAGTTGGGAAACAAAGAGGCATATAAGTGTATGAGCTTGGGCAACAGAGAGCACTTTG 1680  
Qy 1681 CATCTGTCTGCAAGCTCACTCTACATGTAATGTAATGTAATGTAATGTAATGTAATG 1740  
Db 1681 CATCTGTCTGCAAGCTCACTCTACATGTAATGTAATGTAATGTAATGTAATGTAATG 1740  
Qy 1741 CCCAAAGAAACGGGCTACACAGAGTTAGTAAGTCTTTTAAAGAAACCTGGAAGTTAA 1800  
Db 1741 CCCAAAGAAACGGGCTACACAGAGTTAGTAAGTCTTTTAAAGAAACCTGGAAGTTAA 1800  
Qy 1801 TCCGAGATGAAGGCTTTGCAAGTGAATGAATAAGCAAGAGTCTCTTCCCTGCTGAGGATG 1860  
Db 1801 TCCGAGATGAAGGCTTTGCAAGTGAATGAATAAGCAAGAGTCTCTTCCCTGCTGAGGATG 1860  
Qy 1861 AAAACCTGAGGAAAGGGGACTGGAGCCAGTTTCAAGCTGATGATGATGATGATGATGATG 1920  
Db 1861 AAAACCTGAGGAAAGGGGACTGGAGCCAGTTTCAAGCTGATGATGATGATGATGATGATG 1920  
Qy 1921 ATGAAGTCTGCAAGAGGCTCTTAAACCTGTAATGTAATGTAATGTAATGTAATGTAATG 1980  
Db 1921 ATGAAGTCTGCAAGAGGCTCTTAAACCTGTAATGTAATGTAATGTAATGTAATGTAATG 1980  
Qy 1981 CAACAGGATGCAAGAGGACAGATCAAAATTTCCATCATGCAACCCCGGGACTCACGTGT 2040  
Db 1981 CAACAGGATGCAAGAGGACAGATCAAAATTTCCATCATGCAACCCCGGGACTCACGTGT 2040  
Qy 2041 GGCCGCAACAGAGGCCCCAACAATGCAAGCTCCGAATGCACTGGGCTTTGGTGTATCCCA 2100  
Db 2041 GGCCGCAACAGAGGCCCCAACAATGCAAGCTCCGAATGCACTGGGCTTTGGTGTATCCCA 2100  
Qy 2101 AGGAAGGCTCCAGATTTCCATGTCGAAGAGACAGGACCTGGGAGGAGCAAGGTGC 2160  
Db 2101 AGGAAGGCTCCAGATTTCCATGTCGAAGAGACAGGACCTGGGAGGAGCAAGGTGC 2160  
Qy 2161 TCATCTTTGATGATCCTTTGAGCAGAGGATGCGCAGGATGCTCATCTTTCCGGCTGA 2220  
Db 2161 TCATCTTTGATGATCCTTTTGAACAGAGGATGCGCAGGATGCTCATCTTTCCGGCTGA 2220  
Qy 2221 TATTTCATGCTGGATGTTGGCATCCGGAATCTGACCAACAGAGAGACGACGCTTCCAG 2280  
Db 2221 TATTTCATGCTGGATGTTGGCATCCGGAATCTGACCAACAGAGAGACGACGCTTCCAG 2280

Db 2221 TATTTCATGCTGGATGTTGGCATCCGGAATCTGACCAACAGAGAGACGACGCTTCCAG 2280  
Qy 2281 CAATTTAGCATGAATTCATCAAGCTTTGGAAATCTCTGGAGAGA 2324  
Db 2281 CAATTTAGCATGAATTCATCAAGCTTTGGAAATCTCTGGAGAGA 2324  
RESULT 4  
US-09-903-216-3  
; Sequence 3, Application US/09903216  
; Patent No. US20020114811A1  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV2  
; CURRENT APPLICATION NUMBER: US/09/903,216  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-903-216-3  
Query Match 100.0%; Score 2324; DB 9; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CGGACCGTGCATATGGCCAGCGTAAGAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
Db 1 CGGACCGTGCATATGGCCAGCGTAAGAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
Qy 61 GCTCCGGCAGCGGTAGCAGAGTGGCGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 120  
Db 61 GCTCCGGCAGCGGTAGCAGAGTGGCGGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 120  
Qy 121 AGCATGTGAGGACACAGAGATGGAGGAGCGGACTCTCGGAACTTCAATCTTCACGT 180  
Db 121 AGCATGTGAGGACACAGAGATGGAGGAGCGGACTCTCGGAACTTCAATCTTCACGT 180  
Qy 181 GGTTCATGCTGATGTCATTTGCTGGCGTCTGGACATCTGTCGTCTGCTTGTGATC 240  
Db 181 GGTTCATGCTGATGTCATTTGCTGGCGTCTGGACATCTGTCGTCTGCTTGTGATC 240  
Qy 241 TTGTTGACTATGAGGAAGTTCTAGAAAACTAGAAATCTATGATGCTGATGCTGATGAG 300  
Db 241 TTGTTGACTATGAGGAAGTTCTAGAAAACTAGAAATCTATGATGCTGATGCTGATGAG 300  
Qy 301 ATTTTCATGTCATGATGTCACAAAGTTTATTTAGGACTTAAAGAGAGATCTCTTCAGAGC 360  
Db 301 ATTTTCATGTCATGTCACAAAGTTTATTTAGGACTTAAAGAGAGATCTCTTCAGAGC 360  
Qy 361 CAGCAGTCCCGCAGAAAGGCTGAGCCACACACTGAGCCCGGAGGAGAGGTTCTTGTGG 420  
Db 361 CAGCAGTCCCGCAGAAAGGCTGAGCCACACACTGAGCCCGGAGGAGAGGTTCTTGTGG 420  
Qy 421 AGCGAGAAACCCAGAAATATCGAAGTGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG 480  
Db 421 AGCGAGAAACCCAGAAATATCGAAGTGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG 480  
Qy 481 AAATGTTACACCGAGAACATGTTGAGCGAGAGACTTTGCAACAGAGAGATGAGCCACAG 540  
Db 481 AAATGTTACACCGAGAACATGTTGAGCGAGAGACTTTGCAACAGAGAGATGAGCCACAG 540  
Qy 541 GAGAACCAACAGAGAGATGATGATGCTTATGCGGACTGATGATGATGATGATGATGATG 600  
Db 541 GAGAACCAACAGAGAGATGATGATGCTTATGCGGACTGATGATGATGATGATGATGATG 600

601 AGACCTTGGAACTGAAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGAGAGA 660  
Db |||||  
601 AGACCTTGGAACTGAAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGAGAGA 660  
Qy |||||  
661 CAGTTTCACAGACTGTAATCAGGATATGGAAGAGATGATGTTCTGAGCAGGAAATCCAG 720  
Db |||||  
661 CAGTTTCACAGACTGTAATCAGGATATGGAAGAGATGATGTTCTGAGCAGGAAATCCAG 720  
Qy |||||  
721 ATTCCAGTGAAACCACTAGTAGAAGATGAAGATTTGCCACCATGATACAGATGATGTAACAT 780  
Db |||||  
721 ATTCCAGTGAAACCACTAGTAGAAGATGAAGATTTGCCACCATGATACAGATGATGTAACAT 780  
Qy |||||  
781 ACCAAGTCTATGAGGAAACCAAGCAGTATATGAACTCTAGAAAATGAAGGATGAAATCA 840  
Db |||||  
781 ACCAAGTCTATGAGGAAACCAAGCAGTATATGAACTCTAGAAAATGAAGGATGAAATCA 840  
Qy |||||  
841 CAGAAGTAACTGCTCCCTCAGGATATCTCTGAGAGATTTACAGGATATTTGTAAG 900  
Db |||||  
841 CAGAAGTAACTGCTCCCTCAGGATATCTCTGAGAGATTTACAGGATATTTGTAAG 900  
Qy |||||  
901 AAGTAAGCATTTTCTCTGGAAGAACAGCAGGAGTACCAACCAAGAACTTAAAT 1020  
Db |||||  
901 AAGTAAGCATTTTCTCTGGAAGAACAGCAGGAGTACCAACCAAGAACTTAAAT 1020  
Qy |||||  
961 CAGATGATCCAGAACAAAGCMAAGTTTAAAGAAAGAAAGCCTTAAAT 1080  
Db |||||  
961 CAGATGATCCAGAACAAAGCMAAGTTTAAAGAAAGAAAGCCTTAAAT 1080  
Qy |||||  
1021 TTGATAAGCATTTAAAGCTGAACTTTGATGCTGAGAAATACTCCGTAAGGGGAAAA 1080  
Db |||||  
1021 TTGATAAGCATTTAAAGCTGAACTTTGATGCTGAGAAATACTCCGTAAGGGGAAAA 1080  
Qy |||||  
1081 TTGAGAAAGCAGTGAATGATTTAAAGAACTAGTACGCAATACCTCAGAGTCCACGAG 1140  
Db |||||  
1081 TTGAGAAAGCAGTGAATGATTTAAAGAACTAGTACGCAATACCTCAGAGTCCACGAG 1140  
Qy |||||  
1141 CAAGATATGGGAAGCGGAGTGTAGGATGATTTGGCTGAGAAGAGGAGAAATGAGG 1200  
Db |||||  
1141 CAAGATATGGGAAGCGGAGTGTAGGATGATTTGGCTGAGAAGAGGAGAAATGAGG 1200  
Qy |||||  
1201 TGCTACGTGGAGCCATCGAGACCTACCAAGAGTGGCCAGCTCCTGATGTCCTGAG 1260  
Db |||||  
1201 TGCTACGTGGAGCCATCGAGACCTACCAAGAGTGGCCAGCTCCTGATGTCCTGAG 1260  
Qy |||||  
1261 ACTGCTGAAGCTGATTTGAAGCTCGCTCAGACAGGCAACATTTCTAGTCAATGA 1320  
Db |||||  
1261 ACTGCTGAAGCTGATTTGAAGCTCGCTCAGACAGGCAACATTTCTAGTCAATGA 1320  
Qy |||||  
1321 GAGGTTCCCTGCTTACCTCGCAGAGATTTAGTTCAACTATTTCCCAATGATATCTCTTAA 1380  
Db |||||  
1321 GAGGTTCCCTGCTTACCTCGCAGAGATTTAGTTCAACTATTTCCCAATGATATCTCTTAA 1380  
Qy |||||  
1381 AAAATGACCTTGGCTGGGATACCTCTTGATAGGAGATTAAGCAATGCAAGAAAGTTT 1440  
Db |||||  
1381 AAAATGACCTTGGCTGGGATACCTCTTGATAGGAGATTAAGCAATGCAAGAAAGTTT 1440  
Qy |||||  
1441 ATGAGAGTGTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTATGCTTCA 1500  
Db |||||  
1441 ATGAGAGTGTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTATGCTTCA 1500  
Qy |||||  
1501 TCCTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCATATTTAAAGGAAGAAATAGAT 1560  
Db |||||  
1501 TCCTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCATATTTAAAGGAAGAAATAGAT 1560  
Qy |||||  
1561 CCGGAGATCTGCACTGATGAGGATTTTATTTCCACCTGGGGATGCCATGCAGA 1620  
Db |||||  
1561 CCGGAGATCTGCACTGATGAGGATTTTATTTCCACCTGGGGATGCCATGCAGA 1620  
Qy |||||  
1621 GGGTTGGGAACAAAGAGGCATATAAGTGTATGAGCTTGGGCACAGAGAGACATTTG 1680  
Db |||||  
1621 GGGTTGGGAACAAAGAGGCATATAAGTGTATGAGCTTGGGCACAGAGAGACATTTG 1680

1681 CATCTGTGGCAACGCTCACTCTCAATGTGAATGGAAGCAAGCAAGCTTGGTGA 1740  
Db |||||  
1681 CATCTGTGGCAACGCTCACTCTCAATGTGAATGGAAGCAAGCAAGCTTGGTGA 1740  
Qy |||||  
1741 CCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAAGAACTGGAAGTTAA 1800  
Db |||||  
1741 CCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAAGAACTGGAAGTTAA 1800  
Qy |||||  
1801 TCCGAGATGAAGGCTTTCAGTGTGATTAAGCCAAAGGCTCTTCTCTGCTCAGGATG 1860  
Db |||||  
1801 TCCGAGATGAAGGCTTTCAGTGTGATTAAGCCAAAGGCTCTTCTCTGCTCAGGATG 1860  
Qy |||||  
1861 AAAACCTGAGGAAAGGGGCTGAGCCAGTTCACCGCTGTGGCAGCAGAAAGAA 1920  
Db |||||  
1861 AAAACCTGAGGAAAGGGGCTGAGCCAGTTCACCGCTGTGGCAGCAGAAAGAA 1920  
Qy |||||  
1921 ATGAAAATCCCTGCAAGGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGGAGA 1980  
Db |||||  
1921 ATGAAAATCCCTGCAAGGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGGAGA 1980  
Qy |||||  
1981 CAAACGATGCAAGAGAGGACAGATCAAAATATTTCCATCATGCAACCCCGGAGTCAAGTGT 2040  
Db |||||  
1981 CAAACGATGCAAGAGAGGACAGATCAAAATATTTCCATCATGCAACCCCGGAGTCAAGTGT 2040  
Qy |||||  
2041 GGCCGACACAGGCCCCCAAACTGAGGCTCGAATGCACTGGCTTGGTGAATTTCCCA 2100  
Db |||||  
2041 GGCCGACACAGGCCCCCAAACTGAGGCTCGAATGCACTGGCTTGGTGAATTTCCCA 2100  
Qy |||||  
2101 AGCAAGGCTGCAAGATTCGATGTGCCAACGAGACAGGACCTGGGAGGAGCAAGGTGC 2160  
Db |||||  
2101 AGCAAGGCTGCAAGATTCGATGTGCCAACGAGACAGGACCTGGGAGGAGCAAGGTGC 2160  
Qy |||||  
2161 TCATCTTTGATGACTCTTTTGAGCAGAGATGCGAGATGCGAGATGCTCATCTTCCGGCTGA 2220  
Db |||||  
2161 TCATCTTTGATGACTCTTTTGAGCAGAGATGCGAGATGCGAGATGCTCATCTTCCGGCTGA 2220  
Qy |||||  
2221 TATTCATGTTGATGCTGTCGCACTCCGAACTGACACACAGCAGAGACGAGCTTCCAG 2280  
Db |||||  
2221 TATTCATGTTGATGCTGTCGCACTCCGAACTGACACACAGCAGAGACGAGCTTCCAG 2280  
Qy |||||  
2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTCTGGAGAGA 2324  
Db |||||  
2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTCTGGAGAGA 2324

RESULT 5

US-09-903-199-3  
; Sequence 3, Application US/09903199  
; Patent No. US20020122802A1  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV4  
; CURRENT APPLICATION NUMBER: US/09/903,199  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-903-199-3

Query Match 100.0%; Score 2324; DB 9; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCGTGAATGGCCAGCGTGAAGATGCCAAGAGCGCGGCAACAGCAGCAGCGG 60



QY	2221	TATTTCATCGTGGATGTGTGGCATCCGAACTGACACACACAGCAGACGCGCCTTCCAG	2280
Db	2221	TATTTCATCGTGGATGTGTGGCATCCGAACTGACACACACAGCAGACGCGCCTTCCAG	2280
QY	2281	CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA	2324
Db	2281	CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA	2324
RESULT 6			
US-09-903-023-3			
; Sequence 3, Application US/09903023			
; Patent No. US20020146421A1			
; GENERAL INFORMATION:			
; APPLICANT: Wands, Jack R.			
; APPLICANT: de la Monte, Suzanne M.			
; APPLICANT: Ince, Nedim			
; APPLICANT: Carlson, Rolf I.			
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS			
; FILE REFERENCE: 21486-032 DIV1			
; CURRENT APPLICATION NUMBER: US/09/903.023			
; CURRENT FILING DATE: 2001-10-11			
; PRIOR APPLICATION NUMBER: 09/436,184			
; PRIOR FILING DATE: 1999-11-08			
; NUMBER OF SEQ ID NOS: 9			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 3			
; LENGTH: 2324			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-903-023-3			
Query Match 100.0%; Score 2324; DB 9; Length 2324;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	CGGACCGTGCATGCGCCAGCGTAGAATGCCAGAGCAGCGGCAACAGCAGCAGCGCG	60
Db	1	CGGACCGTGCATGCGCCAGCGTAGAATGCCAGAGCAGCGGCAACAGCAGCAGCGCG	60
QY	61	GCTCCGGCAGCGGTAGCAGAGTGGCGGCGCAGCAGCGCCCGGGCCCGGAGAGACAA	120
Db	61	GCTCCGGCAGCGGTAGCAGAGTGGCGGCGCAGCAGCGCCCGGGCCCGGAGAGACAA	120
QY	121	AGCATGGAGGACACAAGAAATGGAGGAAAGGGCGACTCTCGGGAACCTTCATCTTCAGT	180
Db	121	AGCATGGAGGACACAAGAAATGGAGGAAAGGGCGACTCTCGGGAACCTTCATCTTCAGT	180
QY	181	GGTTTATGCTGATTTGATTTGCTGGCGCTCTGGACATCTGTAGCTCTCGTTTGGTTGATC	240
Db	181	GGTTTATGCTGATTTGATTTGCTGGCGCTCTGGACATCTGTAGCTCTCGTTTGGTTGATC	240
QY	241	TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGGTGTAGAG	300
Db	241	TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGGTGTAGAG	300
QY	301	ATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	360
Db	301	ATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	360
QY	361	CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGCGAGTTCTCTGG	420
Db	361	CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGCGAGTTCTCTGG	420
QY	421	AGGCAGAACCCAGATATGCAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC	480
Db	421	AGGCAGAACCCAGATATGCAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC	480
QY	481	AAATGCTACCGCAGACATGTTGAGGAGAGACTTGCACAAAGAAATGGACCCACAG	540
Db	481	AAATGCTACCGCAGACATGTTGAGGAGAGACTTGCACAAAGAAATGGACCCACAG	540

QY	541	GAGAACCAACAAGAGGATGATGAGTTTCTTATGGCGACTGATGATGATGATGATGATGATG	600
Db	541	GAGAACCAACAAGAGGATGATGAGTTTCTTATGGCGACTGATGATGATGATGATGATGATG	600
QY	601	AGACCTCGAACCTGAAAGTATCTCATGAAGAACCGAGCATAGTTACCACTGGAGAGA	660
Db	601	AGACCTCGAACCTGAAAGTATCTCATGAAGAACCGAGCATAGTTACCACTGGAGAGA	660
QY	661	CAGTTTCAAGACCTGTAATCAGGATATGGAAGAGATGATGTCAGCAGGAAATCCAG	720
Db	661	CAGTTTCAAGACCTGTAATCAGGATATGGAAGAGATGATGTCAGCAGGAAATCCAG	720
QY	721	ATTCCAGTGAACCACTAGTAGAAGATGAAGATGCAATGATGATGATGATGATGATGATG	780
Db	721	ATTCCAGTGAACCACTAGTAGAAGATGAAGATGCAATGATGATGATGATGATGATGATG	780
QY	781	ACCAAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA	840
Db	781	ACCAAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA	840
QY	841	CAGAACTAACTGCTCCCTCCTGAGGATTAATCTGTGTAAGATTCACAGGTAATTTAGAG	900
Db	841	CAGAACTAACTGCTCCCTCCTGAGGATTAATCTGTGTAAGATTCACAGGTAATTTAGAG	900
QY	901	AAGTAAGCAATTTTCTGTGGAAGAACAGCAGGAGTACCAACAGAAACAAATAGAA	960
Db	901	AAGTAAGCAATTTTCTGTGGAAGAACAGCAGGAGTACCAACAGAAACAAATAGAA	960
QY	961	CAGATGATCCGAAACAAAGCAAAAGTTAAGAAAGAAAGCTAACTTTTAAATTAAT	1020
Db	961	CAGATGATCCGAAACAAAGCAAAAGTTAAGAAAGAAAGCTAACTTTTAAATTAAT	1020
QY	1021	TTGATAAGCACTATTAAAGCTGAACTTGTGTCAGAAAAAATCCGTAAGAGGGA	1080
Db	1021	TTGATAAGCACTATTAAAGCTGAACTTGTGTCAGAAAAAATCCGTAAGAGGGA	1080
QY	1081	TTGAGGAAGCAGTAATGCAATTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG	1140
Db	1081	TTGAGGAAGCAGTAATGCAATTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG	1140
QY	1141	CAAGATATGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAAAGAGAGTAATGAGG	1200
Db	1141	CAAGATATGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAAAGAGAGTAATGAGG	1200
QY	1201	TGCTACGTGAGCAGCAGCCTACCAAGAGTGGCGCCTACCTGATGTCCTGCGAG	1260
Db	1201	TGCTACGTGAGCAGCAGCCTACCAAGAGTGGCGCCTACCTGATGTCCTGCGAG	1260
QY	1261	ACCTGCTGAAGCTGAGTTTGAAGCGTCTCAGACAGGCAACAAATTTCTAGGTCATATGA	1320
Db	1261	ACCTGCTGAAGCTGAGTTTGAAGCGTCTCAGACAGGCAACAAATTTCTAGGTCATATGA	1320
QY	1321	GAGTTTCCCTGCTTACCTCCTCAGAGATTTAGTTCAACTTTTCCCAATGATATCTTCTAA	1380
Db	1321	GAGTTTCCCTGCTTACCTCCTCAGAGATTTAGTTCAACTTTTCCCAATGATATCTTCTAA	1380
QY	1381	AAATGATCCTTGGCGTGGGATACCTCTTGTAGAGAGATTAATGCAATGCAAGAAAGTTT	1440
Db	1381	AAATGATCCTTGGCGTGGGATACCTCTTGTAGAGAGATTAATGCAATGCAAGAAAGTTT	1440
QY	1441	ATGAAGAGGTGCTGAGTGTGACACCTAATGATGCTTTGCTTAAAGTCAATTTATGGCTTCA	1500
Db	1441	ATGAAGAGGTGCTGAGTGTGACACCTAATGATGCTTTGCTTAAAGTCAATTTATGGCTTCA	1500
QY	1501	TCCTGAAGGCAACAGAAATAATGCTGAGAGATCCCATATTTAAGAGGAGATAGAT	1560
Db	1501	TCCTGAAGGCAACAGAAATAATGCTGAGAGATCCCATATTTAAGAGGAGATAGAT	1560
QY	1561	CGGAGATCTCTGGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG	1620
Db	1561	CGGAGATCTCTGGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG	1620
QY	1621	GGTGGGAAACAAGAGGCAATATAAGTGGTATGAGTTGGGCAACAAGAGGACACTTTG	1680
Db	1621	GGTGGGAAACAAGAGGCAATATAAGTGGTATGAGTTGGGCAACAAGAGGACACTTTG	1680



```
Db 1621 GGGTTGGGAACAAGAGGCAATATAGTGGTATGAGTCTGGGCAAGAGAGGACACTTTG 1680
Qy 1681 CATCTGTCTGGCAACGCTCACTCTACAATGTGAATGGAAGTGAAGCAAGAGGCTTTGGTGGGA 1740
Db 1681 CATCTGTCTGGCAACGCTCACTCTACAATGTGAATGGAAGTGAAGCAAGAGGCTTTGGTGGGA 1740
Qy 1741 CCCCCAAAAAAGGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAAGAAATCGGAAGTTAA 1800
Db 1741 CCCCCAAAAAAGGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAAGAAATCGGAAGTTAA 1800
Qy 1801 TCCGAGATGAAGGCTTTCAGTGTATGATATAAGCCAAAGGCTCTCTTCCTGGCTGAGAGTG 1860
Db 1801 TCCGAGATGAAGGCTTTCAGTGTATGATATAAGCCAAAGGCTCTCTTCCTGGCTGAGAGTG 1860
Qy 1861 AAAAAGCTGAGGGAAGAGGAGCTGGAGCCAGTTACGCTGTGGCAGCAAGGAAGAA 1920
Db 1861 AAAAAGCTGAGGGAAGAGGAGCTGGAGCCAGTTACGCTGTGGCAGCAAGGAAGAA 1920
Qy 1921 ATGAAATGCTTGAAGAGGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGGAGA 1980
Db 1921 ATGAAATGCTTGAAGAGGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGGAGA 1980
Qy 1981 CAACAGGATGCAAGAGGAGGACAGATCAATATATTTCCATCATGACCCCGGAGCTCAGTGT 2040
Db 1981 CAACAGGATGCAAGAGGAGGACAGATCAATATATTTCCATCATGACCCCGGAGCTCAGTGT 2040
Qy 2041 GGCCGCACACAGGGCCCAAACTGCGAGGCTCGGAATGCACTGGGCTTGGTGAATCCCA 2100
Db 2041 GGCCGCACACAGGGCCCAAACTGCGAGGCTCGGAATGCACTGGGCTTGGTGAATCCCA 2100
Qy 2101 AGGAAGCTGCAAGATTCGATGTGCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
Db 2101 AGGAAGCTGCAAGATTCGATGTGCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
Qy 2161 TCATCTTTGATGACTCTTTGAGCAGAGGATGAGCAGGATGAGCAGGATGAGCAGGATGAG 2220
Db 2161 TCATCTTTGATGACTCTTTGAGCAGAGGATGAGCAGGATGAGCAGGATGAGCAGGATGAG 2220
Qy 2221 TATTCATCGTGGATGTGGCATCCGGAGCTGACACACAGCAGCAGCAGCAGCAGCAGCAG 2280
Db 2221 TATTCATCGTGGATGTGGCATCCGGAGCTGACACACAGCAGCAGCAGCAGCAGCAGCAG 2280
Qy 2281 CAATTTAGCATGAATTCAGTCAGCTTGGGAAACTCTGGAGAGA 2324
Db 2281 CAATTTAGCATGAATTCAGTCAGCTTGGGAAACTCTGGAGAGA 2324
```

## RESULT 7

```
US-09-436-184-3
; Sequence 3, Application US/09436184
; Publication No. US20030031670A1
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R.
; APPLICANT: de la Monte, Suzanne M.
; APPLICANT: Ince, Nedim
; APPLICANT: Carlson, Rolf I.
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS
; FILE REFERENCE: R.I. Hosp. Malignant Neoplasms
; CURRENT APPLICATION NUMBER: US/09/436.184
; CURRENT FILING DATE: 1999-11-08
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2324
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-436-184-3
```

```
Query Match 100.0%; Score 2324; DB 10; Length 2324;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 CGGACCGTGCATATGCGCCCGGCTAAGATGCGCAAGAGCAGCGGCAACAGCGAGCGAGCG 60
Db 1 CGGACCGTGCATATGCGCCCGGCTAAGATGCGCAAGAGCAGCGGCAACAGCGAGCGAGCG 60
Qy 61 GCTCCGGCAGCGGTAGCACAGGTGCGGCGAGCAGCGCCCGGGGCGCGAGAGAGACAA 120
Db 61 GCTCCGGCAGCGGTAGCACAGGTGCGGCGAGCAGCGCCCGGGGCGCGAGAGAGACAA 120
Qy 121 AGCATGAGAGACACAAGATGCGAGAAAGCGGAGCTCTCGGGAACCTTCTTCACTT 180
Db 121 AGCATGAGAGACACAAGATGCGAGAAAGCGGAGCTCTCGGGAACCTTCTTCACTT 180
Qy 181 GGTATTATGGTGAATTCGATTCGCGGCTCTGGACATCTGTAGCTGTCTGTTGGTTGATC 240
Db 181 GGTATTATGGTGAATTCGATTCGCGGCTCTGGACATCTGTAGCTGTCTGTTGGTTGATC 240
Qy 241 TTGTTGACTATGAGGAGTCTTAGGAAACTAGGAACTATGATGCTGATGCTGATGCTGAG 300
Db 241 TTGTTGACTATGAGGAGTCTTAGGAAACTAGGAACTATGATGCTGATGCTGATGCTGAG 300
Qy 301 ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Db 301 ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
Qy 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGTTG 420
Db 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGTTG 420
Qy 421 AGGCAGAAACCCCAAGATATCGAAGATGAAGCAAGAAAGCAAAATTCAGTCTCTTCCATG 480
Db 421 AGGCAGAAACCCCAAGATATCGAAGATGAAGCAAGAAAGCAAAATTCAGTCTCTTCCATG 480
Qy 481 AAATGTTACACGAGCAACATGTTGAGGAGAGGAGCTTGCACAAAGAGATGAGCCACAG 540
Db 481 AAATGTTACACGAGCAACATGTTGAGGAGAGGAGCTTGCACAAAGAGATGAGCCACAG 540
Qy 541 GAGAACCAACAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Db 541 GAGAACCAACAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
Qy 601 AGACCTTGAACCTCGAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGAAGAGA 660
Db 601 AGACCTTGAACCTCGAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGAAGAGA 660
Qy 661 CAGTTTACAGACTGTAATCAGGATATGGAAGAGATGATGATGATGATGATGATGATGATG 720
Db 661 CAGTTTACAGACTGTAATCAGGATATGGAAGAGATGATGATGATGATGATGATGATGATG 720
Qy 721 ATTCCAGTGAACCTAGTAGAAGTGAAGATGCAACCATGATACAGATGATGATGATGATG 780
Db 721 ATTCCAGTGAACCTAGTAGAAGTGAAGATGCAACCATGATACAGATGATGATGATGATG 780
Qy 781 ACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA 840
Db 781 ACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA 840
Qy 841 CAGAAGTAACTCTCCCTCGAGGATATCTCTGAGAGATTCACAGGATTCACAGGATTCAG 900
Db 841 CAGAAGTAACTCTCCCTCGAGGATATCTCTGAGAGATTCACAGGATTCACAGGATTCAG 900
Qy 901 AAGTAAAGCAATTTTCTGTTGGAAGAGCAGCAGGAGTACCAAGGAGGAGGAGGAGGAGG 960
Db 901 AAGTAAAGCAATTTTCTGTTGGAAGAGCAGCAGGAGTACCAAGGAGGAGGAGGAGGAGG 960
Qy 961 CAGATGATCCAGAACAAAAAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
Db 961 CAGATGATCCAGAACAAAAAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
Qy 1021 TTGATAAGACTATTTAAAGCTGAACTTGTGCTGAGAGAAAACTCCGTAAGAGGAGGAGG 1080
Db 1021 TTGATAAGACTATTTAAAGCTGAACTTGTGCTGAGAGAAAACTCCGTAAGAGGAGGAGG 1080
Qy 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGCAAAATACCTCCAGAGTCCAG 1140
```



Db 1081 TTGAGGAGCAGTGAATGCAATTTAAAGAACTAGTAGCAAAATACCTCTAGAGTCCAGAG 1140  
Qy 1141 CAAGATATGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAAGTAATGAGG 1200  
Db 1141 CAAGATATGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAAGTAATGAGG 1200  
Qy 1201 TGCTACGTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCCTACCTGATGTCCTCCGAG 1260  
Db 1201 TGCTACGTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCCTACCTGATGTCCTCCGAG 1260  
Qy 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACATTTCTAGGTCATATGA 1320  
Db 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACATTTCTAGGTCATATGA 1320  
Qy 1321 GAGGTTCCCTGCTTACCTCGCAGAGATTAGTTCAACTATTTCCCAATGATACCTCTTAA 1380  
Db 1321 GAGGTTCCCTGCTTACCTCGCAGAGATTAGTTCAACTATTTCCCAATGATACCTCTTAA 1380  
Qy 1381 AAAATGACCTTGGCTGGGATACCTCTGTAGAGATATGACAATGCAAGAAAGTTT 1440  
Db 1381 AAAATGACCTTGGCTGGGATACCTCTGTAGAGATATGACAATGCAAGAAAGTTT 1440  
Qy 1441 ATGAGAGGTGCTGAGTGTGACACCTAATGATGCTTTGCTTAAAGTCCATTTATGGCTTCA 1500  
Db 1441 ATGAGAGGTGCTGAGTGTGACACCTAATGATGCTTTGCTTAAAGTCCATTTATGGCTTCA 1500  
Qy 1501 TCCTGAAGGCAAGAAATTTGCTGAGAGATCCCATATTTAAAGAGGATAGAAAT 1560  
Db 1501 TCCTGAAGGCAAGAAATTTGCTGAGAGATCCCATATTTAAAGAGGATAGAAAT 1560  
Qy 1561 CCGGAGATCTGGCACTGATGATGGAGATTTATTTCCACTGGGGATGCCATGAGA 1620  
Db 1561 CCGGAGATCTGGCACTGATGATGGAGATTTATTTCCACTGGGGATGCCATGAGA 1620  
Qy 1621 GGGTTGGGAAACAAAGGCGCATATAAGTGTATGAGCTTTGGGCACAAAGAGAGACATTTG 1680  
Db 1621 GGGTTGGGAAACAAAGGCGCATATAAGTGTATGAGCTTTGGGCACAAAGAGAGACATTTG 1680  
Qy 1681 CATCTCTGTGGCAACCTCACTCTCAATGTGAATGGAATGGAAGCAAGCCCTTGGTGA 1740  
Db 1681 CATCTCTGTGGCAACCTCACTCTCAATGTGAATGGAATGGAAGCAAGCCCTTGGTGA 1740  
Qy 1741 CCCAAAGAAACGGGCTACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAATTA 1800  
Db 1741 CCCAAAGAAACGGGCTACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAATTA 1800  
Qy 1801 TCCGAGATGAAGGCTTGCAGTGAATGGAATAAGCCAAAGGTCTTCTGCTGAGGATG 1860  
Db 1801 TCCGAGATGAAGGCTTGCAGTGAATGGAATAAGCCAAAGGTCTTCTGCTGAGGATG 1860  
Qy 1861 AAAACCTGAGGAAAAGGGGACTGAGCCGTTTACGCTGTGGCAGCAAGAAAGAA 1920  
Db 1861 AAAACCTGAGGAAAAGGGGACTGAGCCGTTTACGCTGTGGCAGCAAGAAAGAA 1920  
Qy 1921 ATGAAATGCTGCAAGAGAGCTTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGAGA 1980  
Db 1921 ATGAAATGCTGCAAGAGAGCTTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGAGA 1980  
Qy 1981 CAACAGGATGCAAGAGGACAGATCAATATTTCCATCATGCAACCCGGGACTCAGTGT 2040  
Db 1981 CAACAGGATGCAAGAGGACAGATCAATATTTCCATCATGCAACCCGGGACTCAGTGT 2040  
Qy 2041 GGGCGCACAGGGCCCAACAACTGCAAGGCTCCGAATGCACTTGGGCTTGGTGAATTTCCA 2100  
Db 2041 GGGCGCACAGGGCCCAACAACTGCAAGGCTCCGAATGCACTTGGGCTTGGTGAATTTCCA 2100  
Qy 2101 AGGAAGGCTGCAAGATTCGATGTGCCAAACGAGACCGAGGACTGGGAGGAGCAAGTGC 2160  
Db 2101 AGGAAGGCTGCAAGATTCGATGTGCCAAACGAGACCGAGGACTGGGAGGAGCAAGTGC 2160  
Qy 2161 TCATCTTTGATGATCTCTTTGAGCACGAGGTATGGCAGGATGCTCATCTTTCCGGCTGA 2220

Db 2161 TCATCTTTGATGATCTCTTTGAGCACGAGGTATGGCAGGATGCTCATCTTTCCGGCTGA 2220  
Qy 2221 TATTTCATCTGATGCTGTGGCATCCGGAACCTGACACCAACAGCAGAGCGAGCCTTCAG 2280  
Db 2221 TATTTCATCTGATGCTGTGGCATCCGGAACCTGACACCAACAGCAGAGCGAGCCTTCAG 2280  
Qy 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTCTGGAGAGA 2324  
Db 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTCTGGAGAGA 2324

RESULT 8  
US-10-084-817-63  
; Sequence 63, Application US/10084817  
; Publication No. US20030119009A1  
; GENERAL INFORMATION:  
; APPLICANT: Susan Stuart  
; APPLICANT: Jed G. Nuchtern  
; APPLICANT: Sharon E. Pion  
; APPLICANT: Jason M. Shohet  
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION  
; FILE REFERENCE: PA-0046 US  
; CURRENT APPLICATION NUMBER: US/10/084,817  
; CURRENT FILING DATE: 2002-02-25  
; PRIOR APPLICATION NUMBER: 60/270,784  
; PRIOR FILING DATE: 2001-02-23  
; NUMBER OF SEQ ID NOS: 365  
; SOFTWARE: PERL Program  
; SEQ ID NO 63  
; LENGTH: 5358  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20030119009A1 089172.13  
US-10-084-817-63

Query Match 99.8%; Score 2319.2; DB 15; Length 5358;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2321; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGGACCTGCAATGGCCCGCGGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
Db 258 CGGACCTGCAATGGCCCGCGGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 317  
Qy 61 GCTCCGCGCAGCGGTAGCAGCGAGTGGGCGGAGCAGCGCCCGGGCCCGGAGAGACAA 120  
Db 318 GCTCCGCGCAGCGGTAGCAGCGAGTGGGCGGAGCAGCGCCCGGGCCCGGAGAGACAA 377  
Qy 121 AGCATGGAGGACACAAAGATGGGAGGAAAGCGGACTCTCGGAACTTCAATTTCACTG 180  
Db 378 AGCATGGAGGACACAAAGATGGGAGGAAAGCGGACTCTCGGAACTTCAATTTCACTG 437  
Qy 181 GGTATTATGGTGAATTCGATTTGCTGGCGCTGTGGACATCTGTAGCTGTGCTTTGTTGATC 240  
Db 438 GGTATTATGGTGAATTCGATTTGCTGGCGCTGTGGACATCTGTAGCTGTGCTTTGTTGATC 497  
Qy 241 TTGTTGACTATGACGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGCTGATGGAG 300  
Db 498 TTGTTGACTATGACGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGCTGATGGAG 557  
Qy 301 ATTTTGTATGCTGATGATGCCAAAGCTTTTATTTAGGACTTAAAGAGAGATCTATCTTCAG 360  
Db 558 ATTTTGTATGCTGATGATGCCAAAGCTTTTATTTAGGACTTAAAGAGAGATCTATCTTCAG 617  
Qy 361 CAGCAGTCCCGCAGAGAGGCTGAGCAGACACTGAGCCCGGAGCAGGCTTCTCTGG 420  
Db 618 CAGCAGTCCCGCAGAGAGGCTGAGCAGACACTGAGCCCGGAGCAGGCTTCTCTGG 677  
Qy 421 AGGCAGAACCCAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
Db 678 AGGCAGAACCCAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 737

QY 481 AATGTGATACCCAGACATGTTGAGGGAGAGACTTGCAACAAGAGATGAGCCACAG 540  
Db 738 AATGTGTACACGAGAAATGTTGAGGGAGAGACTTGCAACAAGAGATGAGCCACAG 797  
QY 541 GAGAACCAACAAGAGAGATGATGAGTTCTTATGCGAGTATGATGATGATGATGATG 600  
Db 798 GAGAACCAACAAGAGAGATGATGAGTTCTTATGCGAGTATGATGATGATGATGATG 857  
QY 601 AGACCCCTGAGACCTGAGTATCTCATGAGAAACCGAGATAGTTACACGTGGAAGAGA 660  
Db 858 AGACCCCTGAGACCTGAGTATCTCATGAGAAACCGAGATAGTTACACGTGGAAGAGA 917  
QY 661 CAGTTTCACAGACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720  
Db 918 CAGTTTCACAGACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 977  
QY 721 ATTCCAGTGAACCAAGTAGTAGAAGATGGAAGATGGAAGATGGAAGATGGAAGAT 780  
Db 978 ATTCCAGTGAACCAAGTAGTAGAAGATGGAAGATGGAAGATGGAAGATGGAAGAT 1037  
QY 781 ACCAAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAATGGAAGGATAGAAATCA 840  
Db 1038 ACCAAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAATGGAAGGATAGAAATCA 1097  
QY 841 CAGAGTAGTCTGCTCCCTCAGGATATCTGTAAGAGATGTCACAGGTAATGTAGAG 900  
Db 1098 CAGAGTAGTCTGCTCCCTCAGGATATCTGTAAGAGATGTCACAGGTAATGTAGAG 1157  
QY 901 AAGTAGATTTCTGCTGGAAGAACAGCAGAGAGTACACACAGAAACAAATAGAAAAA 960  
Db 1158 AAGTAGATTTCTGCTGGAAGAACAGCAGAGAGTACACACAGAAACAAATAGAAAAA 1217  
QY 961 CAGATGATCCAGAACAAAAAGCAAAAGTTAAGAAAAAGAGCCTTAAATTAAT 1020  
Db 1218 CAGATGATCCAGAACAAAAAGCAAAAGTTAAGAAAAAGAGCCTTAAATTAAT 1277  
QY 1021 TTGATAGACTATTAAAGCTGAACTTGATGCTGCGAAAAAATCCCTGTAAGAGGAGAAAA 1080  
Db 1278 TTGATAGACTATTAAAGCTGAACTTGATGCTGCGAAAAAATCCCTGTAAGAGGAGAAAA 1337  
QY 1081 TTGAGGAAGCGATGATGCAATTAAGAACTAGTACGCAAAATCCCTCAGAGTCCACGAG 1140  
Db 1338 TTGAGGAAGCGATGATGCAATTAAGAACTAGTACGCAAAATCCCTCAGAGTCCACGAG 1397  
QY 1141 CAAGATATGGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGAAATGATGAGG 1200  
Db 1398 CAAGATATGGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGAAATGATGAGG 1457  
QY 1201 TGCTACGTGAGGCATCGAGACCTACCAAGAGGTGGCCAGCTACCTGATGCTCCGCGAG 1260  
Db 1458 TGCTACGTGAGGCATCGAGACCTACCAAGAGGTGGCCAGCTACCTGATGCTCCGCGAG 1517  
QY 1261 ACCTGCTGAGCTGAGTTTGAAGCTGCTCAGACAGGCAACAAATTTCTAGTGCATATGA 1320  
Db 1518 ACCTGCTGAGCTGAGTTTGAAGCTGCTCAGACAGGCAACAAATTTCTAGTGCATATGA 1577  
QY 1321 GAGTTCCCTGCTTACCTGCGAGAGATTAGTTCAATTTCCCAATGATATCTTCTTAA 1380  
Db 1578 GAGTTCCCTGCTTACCTGCGAGAGATTAGTTCAATTTCCCAATGATATCTTCTTAA 1637  
QY 1381 AAAATGACCTTGGCGTGGGATACCTCTTGATAGAGATATGATGATGATGATGATGATG 1440  
Db 1638 AAAATGACCTTGGCGTGGGATACCTCTTGATAGAGATATGATGATGATGATGATGATG 1697  
QY 1441 ATGAGAGGTGCTGAGTGTGACCTAATGATGCTTGTGTAAGTCCATTTATGCTTCA 1500  
Db 1698 ATGAGAGGTGCTGAGTGTGACCTAATGATGCTTGTGTAAGTCCATTTATGCTTCA 1757  
QY 1501 TCCTGAGGCAACGAACAAATTTGCTGAGAGCATGCCATATTTAAAGGAAGAAATAGAAAT 1560  
Db 1758 TCCTGAGGCAACGAACAAATTTGCTGAGAGCATGCCATATTTAAAGGAAGAAATAGAAAT 1817  
QY 1561 CCGGAGATCCTGGCAGTATGATGGAGATTTTATTTCCATCCCTGGGGGATGCCATGCAGA 1620

Db 1818 CCGGAGATCCTGGCAGTATGATGGAGATTTTATTTCCATCCCTGGGGATGCCATGCAGA 1877  
QY 1621 GGGTTGGGAACAAGAGGCAATTAAGTGTATGAGCTTTGGGCAACAAGAGAGGACACTTTG 1680  
Db 1878 GGGTTGGGAACAAGAGGCAATTAAGTGTATGAGCTTTGGGCAACAAGAGAGGACACTTTG 1937  
QY 1681 CATCTGCTGCGCAACGCTCACTCTCAATGTGAATGAACTGAAAGCACACCTTTGTTGA 1740  
Db 1938 CATCTGCTGCGCAACGCTCACTCTCAATGTGAATGAACTGAAAGCACACCTTTGTTGA 1997  
QY 1741 CCCCACAAAGAAACGGGCTACACAGAGTGTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
Db 1998 CCCCACAAAGAAACGGGCTACACAGAGTGTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 2057  
QY 1801 TCCGAGATGAAGGCTTTCAGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1860  
Db 2058 TCCGAGATGAAGGCTTTCAGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 2117  
QY 1861 AAAACCTGAGGGAAAAAGGGGACTCGAGCCAGTTTACGCTGTGCGACAGGAAGAAAGAA 1920  
Db 2118 AAAACCTGAGGGAAAAAGGGGACTCGAGCCAGTTTACGCTGTGCGACAGGAAGAAAGAA 2177  
QY 1921 ATGAAAATGCTCTCAAAAGGAGCTCTCTAAACCTGTACTTACTAGAAAAAGTTCCCGGAGA 1980  
Db 2178 ATGAAAATGCTCTCAAAAGGAGCTCTCTAAACCTGTACTTACTAGAAAAAGTTCCCGGAGA 2237  
QY 1981 CAACAGGATGCAAGAGGAGGACAGATCAATATTTCCATCATGCAACCCGGGACTCAGGTGT 2040  
Db 2238 CAACAGGATGCAAGAGGAGGACAGATCAATATTTCCATCATGCAACCCGGGACTCAGGTGT 2297  
QY 2041 GGGCGCACACAGGGGCCCAAACTGAGGCTTCGAAATGCACTTGGGCTTTGGTGTATCCCA 2100  
Db 2298 GGGCGCACACAGGGGCCCAAACTGAGGCTTCGAAATGCACTTGGGCTTTGGTGTATCCCA 2357  
QY 2101 AGGAAGCTGCAAGATTCGATGTGCAACAGAGACAGAGACCTGGGAGGAAGCAAGGTGC 2160  
Db 2358 AGGAAGCTGCAAGATTCGATGTGCAACAGAGACAGAGACCTGGGAGGAAGCAAGGTGC 2417  
QY 2161 TCATCTTGTATGACTCTCTTTGAGCAGCAGGATGATGGCAGGATGCTCATCTTTCCGGCTGA 2220  
Db 2418 TCATCTTGTATGACTCTCTTTGAGCAGCAGGATGATGGCAGGATGCTCATCTTTCCGGCTGA 2477  
QY 2221 TATTCATCTGAGTGTGTCATCGGATCGGAACTGACACACAGCAGAGACGCGACCTTTCCAG 2280  
Db 2478 TATTCATCTGAGTGTGTCATCGGATCGGAACTGACACACAGCAGAGACGCGACCTTTCCAG 2537  
QY 2281 CAATTTAGCATGAATTCATGCAAGCTTTGGGAAACTCTGGAGAGA 2324  
Db 2538 CAATTTAGCATGAATTCATGCAAGCTTTGGGAAACTCTGGAGAGA 2581

## RESULT 9

US-10-096-534-37  
; Sequence 37, Application US/10096534  
; Publication No. US20030166887A1  
; GENERAL INFORMATION:  
; APPLICANT: The Brigham and Women's Hospital, Inc.  
; APPLICANT: Yates, Karen  
; APPLICANT: Mizuno, Shuichi  
; APPLICANT: Glowacki, Julie  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF SKELETAL DEGENERATION CONDITIONS  
; FILE REFERENCE: B0801/7244/KA/ERP  
; CURRENT APPLICATION NUMBER: US/10/096,534  
; PRIOR FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: US 60/274,980  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 37  
; LENGTH: 2680  
; TYPE: DNA  
; ORGANISM: Homo sapiens

```
US-10-096-534-37
Query Match      40.7%; Score 947; DB 16; Length 2680;
Best Local Similarity 99.5%; Pred. No. 1.3e-214;
Matches 950; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGGACCGTGAATGGCCAGGTAAGATCCAGAGAGCGGCAACAGCAGCAGCAGCG 60
DB 1 CGGACCGTGAATGGCCAGGTAAGATCCAGAGAGCGGCAACAGCAGCAGCAGCG 60

QY 61 GCTCCGCGAGCGGTAGCACAGATGCGGCGAGCAGCAGCAGCAGCAGCAGCA 120
DB 61 GCTCCGCGAGCGGTAGCACAGATGCGGCGAGCAGCAGCAGCAGCAGCAGCA 120

QY 121 AGCATGAGGACACAAGATGGGAGAAAGCGGACTCTCAGGAATCTTCACTT 180
DB 121 AGCATGAGGACACAAGATGGGAGAAAGCGGACTCTCAGGAATCTTCACTT 180

QY 181 GGTATTGCTGATTCGATTCGCGGCTCTGGACATCTGTAGCTGTCTTGTGTTG 240
DB 181 GGTATTGCTGATTCGATTCGCGGCTCTGGACATCTGTAGCTGTCTTGTGTTG 240

QY 241 TTGTTGACTATGAGGAAGTCTTAGGAAAATCTAGGAATCTATGATGCTGATGG 300
DB 241 TTGTTGACTATGAGGAAGTCTTAGGAAAATCTAGGAATCTATGATGCTGATGG 300

QY 301 ATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
DB 301 ATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360

QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCT 420
DB 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCT 420

QY 421 AGGCAGAAACCCAGAAATTCGAAGATGGAAGAAAGAAAGAAAGAAAGAAAG 480
DB 421 AGGCAGAAACCCAGAAATTCGAAGATGGAAGAAAGAAAGAAAGAAAGAAAG 480

QY 481 AATGGTACACGACAACTGATGATGATGATGATGATGATGATGATGATGATG 540
DB 481 AATGGTACACGACAACTGATGATGATGATGATGATGATGATGATGATGATG 540

QY 541 GAGAACCACAACTGATGATGATGATGATGATGATGATGATGATGATGATG 600
DB 541 GAGAACCACAACTGATGATGATGATGATGATGATGATGATGATGATGATG 600

QY 601 AGACCCCTGGAACCTGGAAGTATCTCATGAAGAAACCGAGCATAGTTACCA 660
DB 601 AGACCCCTGGAACCTGGAAGTATCTCATGAAGAAACCGAGCATAGTTACCA 660

QY 661 CAGTTTCAAGACTGTAATCAGGATATGGAAGAGATGATGATGATGATGATG 720
DB 661 CAGTTTCAAGACTGTAATCAGGATATGGAAGAGATGATGATGATGATGATG 720

QY 721 ATTCCAGTGAACGAGTATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 ATTCCAGTGAACGAGTATGATGATGATGATGATGATGATGATGATGATGAT 780

QY 781 ACCAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAAATGAAGGATGA 840
DB 781 ACCAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAAATGAAGGATGA 840

QY 841 CAGAAGTAACTGCTCCCTGAGGATTAATCTGTGAGAGATCTACAGGTAATG 900
DB 841 CAGAAGTAACTGCTCCCTGAGGATTAATCTGTGAGAGATCTACAGGTAATG 900

QY 901 AAGTAAAGCATTTTCTGTTGGAAGACAGCAGGAGTACCACGAAACAAATAG 955
DB 901 AAGTAAAGCATTTTCTGTTGGAAGACAGCAGGAGTACCACGAAACAAATAG 955
```

```
; Sequence 3, Application US/10269909
; Publication No. US20030180747A1
; GENERAL INFORMATION:
; APPLICANT: HRUBAN, RALPH H.
; APPLICANT: ARGANI, PEDRAM
; APPLICANT: IACOBUIO-DONAHUE, CHRISTINE
; APPLICANT: MAITRA, ANIRBAN
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES
; FILE REFERENCE: 58303(71699)
; CURRENT APPLICATION NUMBER: US/10/269,909
; CURRENT FILING DATE: 2003-10-11
; PRIOR APPLICATION NUMBER: 60/328,609
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 60/332,754
; PRIOR FILING DATE: 2001-11-19
; NUMBER OF SEQ ID NOS: 87
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2680
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-269-909-3
```

Query Match 40.7%; Score 947; DB 16; Length 2680;  
Best Local Similarity 99.5%; Pred. No. 1.3e-214;  
Matches 950; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
QY 1 CGGACCGTGAATGGCCAGGTAAGATCCAGAGAGCGGCAACAGCAGCAGCAGCG 60
DB 1 CGGACCGTGAATGGCCAGGTAAGATCCAGAGAGCGGCAACAGCAGCAGCAGCG 60

QY 61 GCTCCGCGAGCGGTAGCACAGATGCGGCGAGCAGCAGCAGCAGCAGCAGCA 120
DB 61 GCTCCGCGAGCGGTAGCACAGATGCGGCGAGCAGCAGCAGCAGCAGCAGCA 120

QY 121 AGCATGAGGACACAAGATGGGAGAAAGCGGACTCTCAGGAATCTTCACTT 180
DB 121 AGCATGAGGACACAAGATGGGAGAAAGCGGACTCTCAGGAATCTTCACTT 180

QY 181 GGTATTGCTGATTCGATTCGCGGCTCTGGACATCTGTAGCTGTCTTGTGTTG 240
DB 181 GGTATTGCTGATTCGATTCGCGGCTCTGGACATCTGTAGCTGTCTTGTGTTG 240

QY 241 TTGTTGACTATGAGGAAGTCTTAGGAAAATCTAGGAATCTATGATGCTGATGG 300
DB 241 TTGTTGACTATGAGGAAGTCTTAGGAAAATCTAGGAATCTATGATGCTGATGG 300

QY 301 ATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
DB 301 ATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360

QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCT 420
DB 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCT 420

QY 421 AGGCAGAAACCCAGAAATTCGAAGATGGAAGAAAGAAAGAAAGAAAGAAAG 480
DB 421 AGGCAGAAACCCAGAAATTCGAAGATGGAAGAAAGAAAGAAAGAAAGAAAG 480

QY 481 AATGGTACACGACAACTGATGATGATGATGATGATGATGATGATGATGATG 540
DB 481 AATGGTACACGACAACTGATGATGATGATGATGATGATGATGATGATGATG 540

QY 541 GAGAACCACAACTGATGATGATGATGATGATGATGATGATGATGATGATG 600
DB 541 GAGAACCACAACTGATGATGATGATGATGATGATGATGATGATGATGATG 600

QY 601 AGACCCCTGGAACCTGGAAGTATCTCATGAAGAAACCGAGCATAGTTACCA 660
DB 601 AGACCCCTGGAACCTGGAAGTATCTCATGAAGAAACCGAGCATAGTTACCA 660

QY 661 CAGTTTCAAGACTGTAATCAGGATATGGAAGAGATGATGATGATGATGATG 720
DB 661 CAGTTTCAAGACTGTAATCAGGATATGGAAGAGATGATGATGATGATGATG 720
```

Db 661 CAGTTTCAACAGACTGTATATCAGGATATGGAAGAGATGATGCTCTGAGCAGGAAATCCAG 720  
Qy 721 ATTCCAGTGAACCAAGTAGTAGAAGATGAAAGATTGCAACATGATACAGATGATGTAACAT 780  
Db 721 ATTCCAGTGAACCAAGTAGTAGAAGATGAAAGATTGCAACATGATACAGATGATGTAACAT 780  
Qy 781 ACCAAGTCTATGAGGAAACAGCAGTATATGAACTCTGAAATGAGGAGGATAGAAATCA 840  
Db 781 ACCAAGTCTATGAGGAAACAGCAGTATATGAACTCTGAAATGAGGAGGATAGAAATCA 840  
Qy 841 CAGAAGTAACTCTCCCTGAGGATATCTGTTGAGAGATTTCACAGTTAATTGTGAAG 900  
Db 841 CAGAAGTAACTCTCCCTGAGGATATCTGTTGAGAGATTTCACAGTTAATTGTGAAG 900  
Qy 901 AAGTAAGCAATTTTCTGTGGAAGAACAGCAGGAGAGTACCACAGAAACAAATAG 955  
Db 901 AAGTAAGCAATTTTCTGTGGAAGAACAGCAGGAGAGTACCACAGATACTTTAAAG 955

## RESULT 11

US-09-736-457-1791  
; Sequence 1791, Application US/09736457  
; Patent No. US20020168637A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Aijun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C15  
; CURRENT APPLICATION NUMBER: US/09/736,457  
; CURRENT FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 1864  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1791  
; LENGTH: 2442  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-736-457-1791

Query Match 35.2%; Score 819; DB 9; Length 2442;  
Best Local Similarity 98.8%; Pred. No. 3.7e-184;  
Matches 825; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
Qy 121 AGCATGAGACACAGAAATGGGAGGAAAGCGGAGCTCTCGGAACTTCATTTTCACGT 180  
Db 5 AGCTTGAAGACACAGAAATGGGAGGAAAGCGGAGCTCTCGGAACTTCATTTTCACGT 64  
Qy 181 GGTATTGTTGATGATGCTGGCGCTCGACATCTGTAGCTGCTGTTGGTTGATC 240  
Db 65 GGTATTGTTGATGATGCTGGCGCTCGACATCTGTAGCTGCTGTTGGTTGATC 124  
Qy 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGATGATGAG 300  
Db 125 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGATGAG 184  
Qy 301 ATTTTGATGTTGATGATGCTAGGAATTTTATTAGGACTTAAAGAGAGTCTACTTCAGGC 360  
Db 185 ATTTTGATGTTGATGATGCTAGGAATTTTATTAGGACTTAAAGAGAGTCTACTTCAGGC 244  
Qy 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACATGAGCCCGCAGGAGAGGTTCTGTGG 420  
Db 245 CAGCAGTCCCGCAGAGAGGCTGAGCCACACATGAGCCCGCAGGAGAGGTTCTGTGG 304  
Qy 421 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAAATTCAGTCCCTTCTCCATG 480

Db 305 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 364  
Qy 481 AAATGGTACACGCGAAGACATGTTGAGGAGAAAGACTTGCACAAAGAAAGATGGACCCACAG 540  
Db 365 AAATGGTACACGCGAAGACATGTTGAGGAGAAAGACTTGCACAAAGAAAGATGGACCCACAG 424  
Qy 541 GAGAACCAACAAAGAGGAGTATGAGTTTCTTATGGCGACTGTATGATGATGATGATGATG 600  
Db 425 GAGAACCAACAAAGAGGAGTATGAGTTTCTTATGGCGACTGTATGATGATGATGATGATG 484  
Qy 601 AGACCTCTGGAACCTGAAAGTATCTCATGAAGAAACCGAGCATAGTTACCCAGTGAAGACA 660  
Db 485 AGACCTCTGGAACCTGAAAGTATCTCATGAAGAAACCGAGCATAGTTACCCAGTGAAGACA 544  
Qy 661 CAGTTTCCAGAGACTGTATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720  
Db 545 CAGTTTCCAGAGACTGTATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 604  
Qy 721 ATTCAGTGAACCCAGTAGTAGAAGATGGAAGATGGAAGATGGAAGATGGAAGATGGAAGAT 780  
Db 605 ATTCAGTGAACCCAGTAGTAGAAGATGGAAGATGGAAGATGGAAGATGGAAGATGGAAGAT 664  
Qy 781 ACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA 840  
Db 665 ACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA 724  
Qy 841 CAGAAGTAACTGCTCCCTGAGGATATCTCTGAGGATATCTCTGAGGATATCTCTGAGGAT 900  
Db 725 CAGAAGTAACTGCTCCCTGAGGATATCTCTGAGGATATCTCTGAGGATATCTCTGAGGAT 784  
Qy 901 AAGTAAGCAATTTTCTGTGGAAGAACAGCAGGAGAGTACCACAGAAACAAATAG 955  
Db 785 AAGTAAGCAATTTTCTGTGGAAGAACAGCAGGAGAGTACCACAGATACTTTAAAG 839

## RESULT 12

US-09-902-941-1791  
; Sequence 1791, Application US/09902941  
; Patent No. US20020172952A1  
; GENERAL INFORMATION:  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Johnson, Jeffrey C.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Mainerakis, Margarita  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: McNabb, Andria  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.478C17  
; CURRENT APPLICATION NUMBER: US/09/902,941  
; CURRENT FILING DATE: 2001-07-10  
; NUMBER OF SEQ ID NOS: 2002  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1791  
; LENGTH: 2442  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-902-941-1791

Query Match 35.2%; Score 819; DB 9; Length 2442;  
Best Local Similarity 98.8%; Pred. No. 3.7e-184;  
Matches 825; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
Qy 121 AGCATGAGACACAGAAATGGGAGGAAAGCGGAGCTCTCGGAACTTCATTTTCACGT 180  
Db 5 AGCTTGAAGACACAGAAATGGGAGGAAAGCGGAGCTCTCGGAACTTCATTTTCACGT 64  
Qy 181 GGTATTGTTGATGATGCTGGCGCTCGACATCTGTAGCTGCTGTTGGTTGATC 240

Db 65 GGTATGGTATGCTGCTGGGCTCTGGCACTCTGATGCTGCTGTTGGTTGATC 124  
QY 241 TTGTTGACTATGAGCAAGTCTTAGGAAACTAGGAATCTATGATGCTGATGCTGATGAG 300  
Db 125 TTGTTGACTATGAGCAAGTCTTAGGAAACTAGGAATCTATGATGCTGATGCTGATGAG 184  
QY 301 ATTTTGTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
Db 185 ATTTTGTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 244  
QY 361 CAGAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGAGGCTTCTGTTGG 420  
Db 245 CAGAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGAGGCTTCTGTTGG 304  
QY 421 AGGCAAGACCCAGAGATATGAGATGAGCAAGCAAAATTCAGTCCCTTCTCCATG 480  
Db 305 AGGCAAGACCCAGAGATATGAGATGAGCAAGCAAAATTCAGTCCCTTCTCCATG 364  
QY 481 AAATGGTACGCGAGAGAGGCTGAGGAGAGAGGCTTGGCAAGAGAGATGAGCCACAG 540  
Db 365 AAATGGTACGCGAGAGAGGCTTGGCAAGAGAGAGGCTTGGCAAGAGAGATGAGCCACAG 424  
QY 541 GAGAACCAACAAGAGATGATGAGTCTTATGCGGAGCTGATGATGATGATGATGATGATG 600  
Db 425 GAGAACCAACAAGAGATGATGAGTCTTATGCGGAGCTGATGATGATGATGATGATGATG 484  
QY 601 AGACCTGGAACTTGAAGTATCTCATGAGAAACCGAGAGATGATGATGATGATGATGATG 660  
Db 485 AGACCTGGAACTTGAAGTATCTCATGAGAAACCGAGAGATGATGATGATGATGATGATG 544  
QY 661 CAGTTTCAAGACTGTAATCAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 720  
Db 545 CAGTTTCAAGACTGTAATCAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 604  
QY 721 ATTCCAGTGAACCTGAGTATGAGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 780  
Db 605 ATTCCAGTGAACCTGAGTATGAGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 664  
QY 781 ACCAAGTCTATGAGAAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 840  
Db 665 ACCAAGTCTATGAGAAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 724  
QY 841 CAGAGTAACTGCTCCCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
Db 725 CAGAGTAACTGCTCCCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 784  
QY 901 AAGTAAGCAATTTTCTGTTGGAAGAACAGAGAGTACCCAGAAACAAATAG 955  
Db 785 AAGTAAGCAATTTTCTGTTGGAAGAACAGAGAGTACCCAGAAACAAATAG 839

RESULT 13  
US-09-849-626-1791  
; Sequence 1791, Application US/09849626  
; Publication No. US20020197669A1  
; GENERAL INFORMATION:  
; APPLICANT: Bangor, Chaitanya  
; APPLICANT: Fanger, Gary  
; APPLICANT: Wang, Aijun  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Switzer, Anne  
; APPLICANT: McNeill, Patricia  
; APPLICANT: Clapper, Jonathan  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 21021.478C16  
; CURRENT APPLICATION NUMBER: US/09/849, 626  
; CURRENT FILING DATE: 2001-05-03  
; NUMBER OF SEQ ID NOS: 1926  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1791  
; LENGTH: 2442

; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-849-626-1791  
Query Match 35.2%; Score 819; DB 9; Length 2442;  
Best Local Similarity 98.8%; Pred. No. 3.7e-184; Indels 0; Gaps 0;  
Matches 825; Conservative 0; Mismatches 10;  
QY 121 AGCATGGAGGACACAAGATGGGAGGAAAGCGGACTCTCGGAACTTCTTCTTCTTCTTCTTCTT 180  
Db 5 ACCTTGAAGGACACAAGATGGGAGGAAAGCGGACTCTCAGGAATCTTCTTCTTCTTCTTCTTCTT 64  
QY 181 GGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240  
Db 65 GGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 124  
QY 241 TTGTTGACTATGAGGAGGATCTAGGAAACTTAGGAACTTAGGAACTTAGGAACTTAGGAACT 300  
Db 125 TTGTTGACTATGAGGAGGATCTAGGAAACTTAGGAACTTAGGAACTTAGGAACTTAGGAACT 184  
QY 301 ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
Db 185 ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 244  
QY 361 CAGAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGAGGCTTCTGTTGG 420  
Db 245 CAGAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGAGGCTTCTGTTGG 304  
QY 421 AGGCAAGACCCAGAGATATGAGATGAGCAAGCAAAATTCAGTCCCTTCTCCATG 480  
Db 305 AGGCAAGACCCAGAGATATGAGATGAGCAAGCAAAATTCAGTCCCTTCTCCATG 364  
QY 481 AAATGGTACGCGAGAGAGGCTGAGGAGAGAGACTTGGCAAGAGAGATGAGCCACAG 540  
Db 365 AAATGGTACGCGAGAGAGGCTTGGCAAGAGAGAGGCTTGGCAAGAGAGATGAGCCACAG 424  
QY 541 GAGAACCAACAAGAGATGATGAGTCTTATGCGGAGCTGATGATGATGATGATGATGATGATG 600  
Db 425 GAGAACCAACAAGAGATGATGAGTCTTATGCGGAGCTGATGATGATGATGATGATGATGATG 484  
QY 601 AGACCTGGAACTTGAAGTATCTCATGAGAAACCGAGAGATGATGATGATGATGATGATGATG 660  
Db 485 AGACCTGGAACTTGAAGTATCTCATGAGAAACCGAGAGATGATGATGATGATGATGATGATG 544  
QY 661 CAGTTTCAAGACTGTAATCAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 720  
Db 545 CAGTTTCAAGACTGTAATCAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 604  
QY 721 ATTCCAGTGAACCTGAGTATGAGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 780  
Db 605 ATTCCAGTGAACCTGAGTATGAGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 664  
QY 781 ACCAAGTCTATGAGAAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 840  
Db 665 ACCAAGTCTATGAGAAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 724  
QY 841 CAGAGTAACTGCTCCCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
Db 725 CAGAGTAACTGCTCCCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 784  
QY 901 AAGTAAGCAATTTTCTGTTGGAAGAACAGAGAGTACCCAGAAACAAATAG 955  
Db 785 AAGTAAGCAATTTTCTGTTGGAAGAACAGAGAGTACCCAGAAACAAATAG 839

RESULT 14  
US-10-017-754-1791  
; Sequence 1791, Application US/10017754  
; Publication No. US20030054363A1  
; GENERAL INFORMATION:  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Watanabe, Yoshihiro

```
APPLICANT: Johnson, Jeffrey C.
APPLICANT: Retter, Marc W.
APPLICANT: Marnerakis, Margarita
APPLICANT: Carter, Darrick
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Bangur, Chaitanya S.
APPLICANT: McNabb, Andria
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.478C18
CURRENT APPLICATION NUMBER: US/10/017,754
NUMBER OF SEQ ID NOS: 2004
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1791
LENGTH: 2442
TYPE: DNA
ORGANISM: Homo sapiens
US-10-017-754-1791

Query Match      35.2%; Score 819; DB 14; Length 2442;
Best Local Similarity 98.8%; Pred. No. 3.7e-184;
Matches 825; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 121 AGCATGGAGGACACAAAGTGGAGGAAAGCGGAGCTCTCGGAACTTCACTTTCACGT 180
DB 5 AGCTTGAAGGACACAAAGTGGAGGAAAGCGGAGCTCTCGGAACTTCACTTTCACGT 64
QY 181 GGTATTATGCTGATTCATTTGCTGGCGCTCTGGACATCTAGCTGCTGTTGGTTGATC 240
DB 65 GGTATTATGCTGATTCATTTGCTGGCGCTCTGGACATCTAGCTGCTGTTGGTTGATC 124
QY 241 TTGTTGACTATGAGGAATCTTAGAAAACCTAGAAATCTATGCTGCTGATGGAG 300
DB 125 TTGTTGACTATGAGGAATCTTAGAAAACCTAGAAATCTATGCTGCTGATGGAG 184
QY 301 ATTTTGATGTGGATGATGCTCAAGTCTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 360
DB 185 ATTTTGATGTGGATGATGCTCAAGTCTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 244
QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGCTCTCTGG 420
DB 245 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGCTCTCTGG 304
QY 421 AGACCTTGGAACTGAGTATCTATGAAGAAACCGGAGCATGTTACCACTGGAAGAGA 660
DB 485 AGACCTTGGAACTGAGTATCTATGAAGAAACCGGAGCATGTTACCACTGGAAGAGA 544
QY 661 CAGTTTCAAGACTGTATCAGGATATGGAAGAGATGATGCTGAGGAGGAAATCCAG 720
DB 545 CAGTTTCAAGACTGTATCAGGATATGGAAGAGATGATGCTGAGGAGGAAATCCAG 604
QY 721 ATTCCAGTGAACCACTAGTAGAAGATGGAAGATGTCACCATCATCATGATGATCAT 780
DB 605 ATTCCAGTGAACCACTAGTAGAAGATGGAAGATGTCACCATCATCATGATGATCAT 664
QY 781 ACCAAGCTTATGAGGAAACAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA 840
DB 665 ACCAAGCTTATGAGGAAACAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA 724

APPLICANT: Watanabe, Yoshihiro
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Sleath, Paul R.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.478C19
CURRENT APPLICATION NUMBER: US/10/113,872
NUMBER OF SEQ ID NOS: 2011
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1791
LENGTH: 2442
TYPE: DNA
ORGANISM: Homo sapiens
US-10-113-872-1791

Query Match      35.2%; Score 819; DB 16; Length 2442;
Best Local Similarity 98.8%; Pred. No. 3.7e-184;
Matches 825; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 121 AGCATGGAGGACACAAAGTGGAGGAAAGCGGAGCTCTCGGAACTTCACTTTCACGT 180
DB 5 AGCTTGAAGGACACAAAGTGGAGGAAAGCGGAGCTCTCGGAACTTCACTTTCACGT 64
QY 181 GGTATTATGCTGATTCATTTGCTGGCGCTCTGGACATCTAGCTGCTGTTGGTTGATC 240
DB 65 GGTATTATGCTGATTCATTTGCTGGCGCTCTGGACATCTAGCTGCTGTTGGTTGATC 124
QY 241 TTGTTGACTATGAGGAATCTTAGAAAACCTAGAAATCTATGCTGCTGATGGAG 300
DB 125 TTGTTGACTATGAGGAATCTTAGAAAACCTAGAAATCTATGCTGCTGATGGAG 184
QY 301 ATTTTGATGTGGATGATGCTCAAGTCTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 360
DB 185 ATTTTGATGTGGATGATGCTCAAGTCTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 244
QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGCTCTCTGG 420
DB 245 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGCTCTCTGG 304
QY 421 AGACCTTGGAACTGAGTATCTATGAAGAAACCGGAGCATGTTACCACTGGAAGAGA 660
DB 485 AGACCTTGGAACTGAGTATCTATGAAGAAACCGGAGCATGTTACCACTGGAAGAGA 544
QY 661 CAGTTTCAAGACTGTATCAGGATATGGAAGAGATGATGCTGAGGAGGAAATCCAG 720
DB 545 CAGTTTCAAGACTGTATCAGGATATGGAAGAGATGATGCTGAGGAGGAAATCCAG 604
QY 721 ATTCCAGTGAACCACTAGTAGAAGATGGAAGATGTCACCATCATCATGATGATCAT 780
DB 605 ATTCCAGTGAACCACTAGTAGAAGATGGAAGATGTCACCATCATCATGATGATCAT 664
QY 781 ACCAAGCTTATGAGGAAACAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA 840
DB 665 ACCAAGCTTATGAGGAAACAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA 724
```

Mon Mar 28 06:06:59 2005

Qy	661	CAGTTTCAAGACTGTAAATCAGGATATGGAAGAGATGATGTCTGAGCGGAAAATCCAG	720
Db	545		604
Qy	721	ATTCCAGTGAAACCGTAGTAGAAGATCAAGATTGCACCATGATACAGATGATGTAAACAT	780
Db	605		664
Qy	781	ACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA	840
Db	665		724
Qy	841	CAGAAAGTAACTGTCTCCCTGAGGATTAATCCTGTAGAGATTCCACAGGTAATTGTAGAAG	900
Db	725		784
Qy	901	AAGTAAGCAATTTTCTGTGGAAGAACAGCAGGGAAGTACCACAGAAACAAATAG	955
Db	785		839

Search completed: March 25, 2005, 17:33:28  
Job time : 1237 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 08:06:08 ; Search time 408 Seconds  
(without alignments)  
9320.361 Million cell updates/sec

Title: US-09-436-184-3

Perfect score: 2324  
Sequence: 1 cggaccgtgaatggccag.....ctgggaactctggagaga 2324

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2324	100.0	2324	4	US-09-903-248-3
2	2324	100.0	2324	4	US-09-903-199-3
3	2324	100.0	2324	4	US-09-903-216-3
4	2324	100.0	2324	4	US-09-903-083-3
5	2324	100.0	2324	4	US-09-853-604-3
6	823.8	35.4	2442	3	US-09-040-485-1
7	819	35.2	2442	4	US-09-702-705-1791
8	819	35.2	2442	4	US-09-736-457-1791
9	819	35.2	2442	4	US-09-671-325-1791
10	539.8	35.2	2442	4	US-09-659-521-1
11	539.8	35.2	2442	4	US-09-949-016-15189
12	153.2	6.6	74730	4	US-09-949-016-15189
13	88.8	3.8	5746	4	US-09-949-016-15189
14	79	3.4	601	4	US-09-949-016-15189
15	77.2	3.3	7218	1	US-08-232-463-14
16	74.6	3.2	3095	6	5231168-1
17	74.6	3.2	3095	6	5231168-1
18	74.4	3.2	601	4	US-09-949-016-122387
19	67.4	2.9	3337	1	US-08-072-610-1
20	67.4	2.9	3337	2	US-08-719-822B-1
21	67.4	2.9	3337	3	US-09-092-458-1
22	67.4	2.9	3337	4	US-08-719-821C-1
23	62	2.7	2208	4	US-09-949-016-236
24	60	2.6	150394	4	US-09-949-016-13042
25	57	2.5	4766	5	PCT-US93-07261-10
26	55.6	2.4	9636	1	US-08-323-170B-1
27	55.6	2.4	9636	3	US-08-954-441-1

US-09-903-248-3  
; Sequence 3, Application US/09903248  
; Patent No. 6783758  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIVS  
; CURRENT APPLICATION NUMBER: US/09/903,248  
; PRIORITY FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-903-248-3

## ALIGNMENTS

### RESULT 1

28	54.8	2.4	696	3	US-09-461-697-193	Sequence 193, App
29	54.8	2.4	699	3	US-09-461-697-191	Sequence 191, App
30	54.8	2.4	717	3	US-09-461-697-189	Sequence 189, App
31	54.8	2.4	774	3	US-09-461-697-187	Sequence 187, App
32	54.8	2.4	819	3	US-09-461-697-185	Sequence 185, App
33	54.8	2.4	1669	3	US-09-461-697-184	Sequence 184, App
34	54.8	2.4	211049	4	US-09-949-016-15770	Sequence 15770, A
35	53.4	2.3	3211	2	US-08-574-959A-8	Sequence 8, Appli
36	53.4	2.3	3211	2	US-09-357-014-8	Sequence 6, Appli
37	53.4	2.3	3901	2	US-08-574-959A-6	Sequence 6, Appli
38	53.4	2.3	3901	2	US-09-357-014-6	Sequence 8, Appli
39	52.8	2.3	614	4	US-09-902-540-1318	Sequence 6, Appli
40	52.8	2.3	1671	4	US-09-248-796A-4970	Sequence 1318, Ap
41	52.4	2.3	601	4	US-09-949-016-122345	Sequence 122345, A
42	52.4	2.3	5361	3	US-08-973-462-2	Sequence 2, Appli
43	52.4	2.3	6152	3	US-08-973-462-1	Sequence 1, Appli
44	50.4	2.2	699	4	US-09-248-796A-9722	Sequence 9722, Ap
45	50.4	2.2	2223	1	US-08-257-073-4	Sequence 4, Appli

Query Match 100.0%; Score 2324; DB 4; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CGCACCGTGCATGTCGCGCCAGCGTAAGATCCCAAGAGAGCGGCAACAGCAGCAGCGCG	60
Db	1	CGCACCGTGCATGTCGCGCCAGCGTAAGATCCCAAGAGAGCGGCAACAGCAGCAGCGCG	60
Qy	61	GCTCCGGAGCGGTAGCAGCGTCCGCGGAGCAGCGCCCGGCGCCGAGAGAGACAA	120
Db	61	GCTCCGGAGCGGTAGCAGCGTCCGCGGAGCAGCGCCCGGCGCCGAGAGAGACAA	120
Qy	121	AGCATGGAGGACACAAAGATGGGAGGAGCGGAGCTCTCGGGAACCTTCTTTCACGT	180
Db	121	AGCATGGAGGACACAAAGATGGGAGGAGCGGAGCTCTCGGGAACCTTCTTTCACGT	180
Qy	181	GGTTTATGGTATGTCGCGGCTCTGGACATCTGTAGCTGTCTGTTGTTGTTGATC	240
Db	181	GGTTTATGGTATGTCGCGGCTCTGGACATCTGTAGCTGTCTGTTGTTGTTGATC	240
Qy	241	TGTTGATCTAGGAGAGTCTAGGAACTAGGATCTATGATGCTGATGGTATGGAG	300
Db	241	TGTTGATCTAGGAGAGTCTAGGAACTAGGATCTATGATGCTGATGGTATGGAG	300
Qy	301	ATTTGATGGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC	360
Db	301	ATTTGATGGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC	360



; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-903-199-3

Query Match 100.0%; Score 2324; DB 4; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCGTGCATTCGCGCCAGCGTAAGAATGCCAAGAGCAGCGCAACAGCAGCAGCGG 60  
DB 1 CGGACCGTGCATTCGCGCCAGCGTAAGAATGCCAAGAGCAGCGCAACAGCAGCAGCGG 60  
QY 61 GCTCCGCGAGCGGTAGCACAGTGCAGGAGCAGCAGCGCCCGGGCCCGGAGAGACAA 120  
DB 61 GCTCCGCGAGCGGTAGCACAGTGCAGGAGCAGCAGCGCCCGGGCCCGGAGAGACAA 120  
QY 121 AGCATGGAGGACACAAATGGAGGAAAGCGGACTCTCGGAACTTCAATCTTCACGT 180  
DB 121 AGCATGGAGGACACAAATGGAGGAAAGCGGACTCTCGGAACTTCAATCTTCACGT 180  
QY 181 GGTATTATGGTGAATTCGCGCGTCTGGACATCTGTAGCTGTGTTGGTTTGTATC 240  
DB 181 GGTATTATGGTGAATTCGCGCGTCTGGACATCTGTAGCTGTGTTGGTTTGTATC 240  
QY 241 TTGTTGACTATGAGGAGTCTAGGAACTAGGAACTATGATGCTGAGTGGAGGAG 300  
DB 241 TTGTTGACTATGAGGAGTCTAGGAACTAGGAACTATGATGCTGAGTGGAGGAG 300  
QY 301 ATTTGATGGATGATGCCAAAGTTTATAGGACTTAAAGAGAGATCTACTTCAGAGC 360  
DB 301 ATTTGATGGATGATGCCAAAGTTTATAGGACTTAAAGAGAGATCTACTTCAGAGC 360  
QY 361 CAGCAGTCCCGCCAGAGAGGTGAGCCACACTGAGCCCGGAGGAGGTTCTGTGG 420  
DB 361 CAGCAGTCCCGCCAGAGAGGTGAGCCACACTGAGCCCGGAGGAGGTTCTGTGG 420  
QY 421 AGGCAGAACCCAGAAATATCGAAGATGAAGCAAGAGACAAATTCAGTCCCTTCCATG 480  
DB 421 AGGCAGAACCCAGAAATATCGAAGATGAAGCAAGAGACAAATTCAGTCCCTTCCATG 480  
QY 481 AAATGGTACACGCAAGACATGTTGAGGAGAGAGACTTGCACAAAGAGAGATGGACCCACAG 540  
DB 481 AAATGGTACACGCAAGACATGTTGAGGAGAGAGACTTGCACAAAGAGAGATGGACCCACAG 540  
QY 541 GAGAACCCACCAAGAGGATGATGATTTCTTATGGGAGTGTATGATGATGATGATTTG 600  
DB 541 GAGAACCCACCAAGAGGATGATGATTTCTTATGGGAGTGTATGATGATGATGATTTG 600  
QY 601 AGACCTCGGAACCTGAAGTATCTCATGAAGAAACCGACATAGTTACACGCTGGAAGAGA 660  
DB 601 AGACCTCGGAACCTGAAGTATCTCATGAAGAAACCGACATAGTTACACGCTGGAAGAGA 660  
QY 661 CAGTTTCACAGACTCTAATCAGGATATGAGAGAGATGATGCTGAGCAGGAGAAATCCAG 720  
DB 661 CAGTTTCACAGACTCTAATCAGGATATGAGAGAGATGATGCTGAGCAGGAGAAATCCAG 720  
QY 721 ATTCAGTGAACCAAGTATGAGAGATGAAGATGACCATGATACAGATGATGATGATGAT 780  
DB 721 ATTCAGTGAACCAAGTATGAGAGATGAAGATGACCATGATACAGATGATGATGATGAT 780  
QY 781 ACCAAGTCTATGAGGAAACAGCAGATATGAACTCTAGAAAATGAAGGAGATAGAAATCA 840  
DB 781 ACCAAGTCTATGAGGAAACAGCAGATATGAACTCTAGAAAATGAAGGAGATAGAAATCA 840  
QY 841 CAGAAGTAACTGCTCCCTCAGGATATCTGTAGAGATTCACAGGTAATTTGTAGAG 900  
DB 841 CAGAAGTAACTGCTCCCTCAGGATATCTGTAGAGATTCACAGGTAATTTGTAGAG 900

DB 841 CAGAAGTAACTGCTCCCTCAGGATATCTGTAGAGATTCACAGGTAATTTGTAGAG 900  
QY 901 AAGTAAGCATTTTCTCTGTGGAAGAAACAGCAGGAAGTACCAACAGAAACAAATAGAAAAA 960  
DB 901 AAGTAAGCATTTTCTCTGTGGAAGAAACAGCAGGAAGTACCAACAGAAACAAATAGAAAAA 960  
QY 961 CAGATGATCCAGAACAAAAAGCTTAAAGAAAGAGAGCTTAAATTAAT 1020  
DB 961 CAGATGATCCAGAACAAAAAGCTTAAAGAAAGAGAGCTTAAATTAAT 1020  
QY 1021 TTGATAAGCATTTTAAAGCTGAACTTGTGCTGCAGAAAAAATCCGCTAAAGGGGAAAAA 1080  
DB 1021 TTGATAAGCATTTTAAAGCTGAACTTGTGCTGCAGAAAAAATCCGCTAAAGGGGAAAAA 1080  
QY 1081 TTGAGGAGCAGTGAATGCAATTTAAAGACTAGTACGCAATACCTCAGAGTCCAGGAG 1140  
DB 1081 TTGAGGAGCAGTGAATGCAATTTAAAGAACTAGTACGCAATACCTCAGAGTCCAGGAG 1140  
QY 1141 CAGATATGGAGAGCGCAGTGTGAGATGATTTGGCTGAGAGAGAGAGTAATGAGG 1200  
DB 1141 CAGATATGGAGAGCGCAGTGTGAGATGATTTGGCTGAGAGAGAGAGTAATGAGG 1200  
QY 1201 TGCTACGTGGAGCCATCGAGACCTACCAAGAGGTGCGCAGCTACCTGATGTCCTGCAG 1260  
DB 1201 TGCTACGTGGAGCCATCGAGACCTACCAAGAGGTGCGCAGCTACCTGATGTCCTGCAG 1260  
QY 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTCTCAGACAGGCAACAATTTCTAGGTCAATAGA 1320  
DB 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTCTCAGACAGGCAACAATTTCTAGGTCAATAGA 1320  
QY 1321 GAGGTTCCTGCTTACCTGCGAGAGATTTAGTTCAACTATTTCCCAATGATCTCTCTTAA 1380  
DB 1321 GAGGTTCCTGCTTACCTGCGAGAGATTTAGTTCAACTATTTCCCAATGATCTCTCTTAA 1380  
QY 1381 AAAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATTAATGCAATGCAAGAAAGTTT 1440  
DB 1381 AAAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATTAATGCAATGCAAGAAAGTTT 1440  
QY 1441 ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGGCTTGTGCTTAAAGTCCATTTAGGCTTCA 1500  
DB 1441 ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGGCTTGTGCTTAAAGTCCATTTAGGCTTCA 1500  
QY 1501 TCCTGAAGGACACAGAAACAAATTTGCTGAGAGCATCCCATTTTAAAGAGAGGATAGAT 1560  
DB 1501 TCCTGAAGGACACAGAAACAAATTTGCTGAGAGCATCCCATTTTAAAGAGAGGATAGAT 1560  
QY 1561 CCGGAGATCTCGGACCTGATGATGGGAGATTTTATTTCCACCTGGGGGATGCCATGCGAGA 1620  
DB 1561 CCGGAGATCTCGGACCTGATGATGGGAGATTTTATTTCCACCTGGGGGATGCCATGCGAGA 1620  
QY 1621 GGGTGGGAAACAAAGAGGCGATATAAGTGTGATGAGCTTTGGGCAACAGAGAGGACACTTTG 1680  
DB 1621 GGGTGGGAAACAAAGAGGCGATATAAGTGTGATGAGCTTTGGGCAACAGAGAGGACACTTTG 1680  
QY 1681 CATCTGTCTGCAACCCCTCACTACATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 1740  
DB 1681 CATCTGTCTGCAACCCCTCACTACATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 1740  
QY 1741 CCCCAAGAGAGAGGCGTACACAGAGTGTAGTAAAGTCTTTTGAAGAAACCTGGAAGTTAA 1800  
DB 1741 CCCCAAGAGAGAGGCGTACACAGAGTGTAGTAAAGTCTTTTGAAGAAACCTGGAAGTTAA 1800  
QY 1801 TCCGAGATGAAGCCCTTGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1860  
DB 1801 TCCGAGATGAAGCCCTTGCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1860  
QY 1861 AAAACCTGAGGAGGAGGAGCTGGAGCCAGTTTCAAGCTGCGCAGAGAGAGAGAGAGAA 1920  
DB 1861 AAAACCTGAGGAGGAGGAGCTGGAGCCAGTTTCAAGCTGCGCAGAGAGAGAGAGAGAA 1920  
QY 1921 ATGAAAATGCTGCAAGAGGCTCTTAAACCTGTA 1980  
DB 1921 ATGAAAATGCTGCAAGAGGCTCTTAAACCTGTA 1980

QY 1981 CAACAGGATGAGAGGAGGACAGATCAAAATATTCATATGCAACCCCGGACTCACGTGT 2040  
Db 1981 CAACAGGATGAGAGGAGGACAGATCAAAATATTCATATGCAACCCCGGACTCACGTGT 2040  
QY 2041 GGCGGCACACAGGCCCCCAAACTGCGAGGTCCGAATGCAACCTGGGCTTGGTATCCCA 2100  
Db 2041 GGCGGCACACAGGCCCCCAAACTGCGAGGTCCGAATGCAACCTGGGCTTGGTATCCCA 2100  
QY 2101 AGAAGGTCGAAGATTCGATGTGCCAACGACGACGACGACCTGGGAGGAAGGCAAGTGC 2160  
Db 2101 AGAAGGTCGAAGATTCGATGTGCCAACGACGACGACGACCTGGGAGGAAGGCAAGTGC 2160  
QY 2161 TCATCTTTGATGACTCTCTTTCAGCACGAGGTATGCGCAGGATGCTCATCTTTCGGGCTGA 2220  
Db 2161 TCATCTTTGATGACTCTCTTTCAGCACGAGGTATGCGCAGGATGCTCATCTTTCGGGCTGA 2220  
QY 2221 TATTCATGCTGGATGTGTGGCATCCGGAATGACACACAGCAGACGACGACCTTCAG 2280  
Db 2221 TATTCATGCTGGATGTGTGGCATCCGGAATGACACACAGCAGACGACGACCTTCAG 2280  
QY 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2324  
Db 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2324

RESULT 3  
US-09-903-216-3  
; Sequence 3, Application US/09903216  
; Patent No. 6812206  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV2  
; CURRENT APPLICATION NUMBER: US/09/903,216  
; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-903-216-3

Query Match 100.0%; Score 2324; DB 4; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCGTGCAATGCGCCAGCGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
Db 1 CGGACCGTGCAATGCGCCAGCGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
QY 61 GCTCCGCGAGCGGTAGCAGGTGCGGCGCAGCAGCAGCGCCCGGCGCCGAGAGACAA 120  
Db 61 GCTCCGCGAGCGGTAGCAGGTGCGGCGCAGCAGCAGCGCCCGGCGCCGAGAGACAA 120  
QY 121 AGCATGGAGGACACAAGAAATGGAGAAAGCGGAACTCTCGGGAATCTTCATCTTCAGGT 180  
Db 121 AGCATGGAGGACACAAGAAATGGAGAAAGCGGAACTCTCGGGAATCTTCATCTTCAGGT 180  
QY 181 GGTTTATGTTGATGATGCTGGGCGTGTGACATCTGTAGCTGTCTGTTGGTTGATC 240  
Db 181 GGTTTATGTTGATGATGCTGGGCGTGTGACATCTGTAGCTGTCTGTTGGTTGATC 240  
QY 241 TTGTTGACTATCAGGAAGTTCTAGGAATACTAGGAATCTATGATGCTGATGGTATGGAG 300  
Db 241 TTGTTGACTATCAGGAAGTTCTAGGAATACTAGGAATCTATGATGCTGATGGTATGGAG 300

QY 301 ATTTGATGTGGATGATGCAAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 360  
Db 301 ATTTGATGTGGATGATGCAAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 360  
QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCCTGTGG 420  
Db 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCCTGTGG 420  
QY 421 AGCGAAGACCCCAAGATATCGAAGATGAAGCAAAAGACAAATTCAGTCCCTTCTCCATG 480  
Db 421 AGCGAAGACCCCAAGATATCGAAGATGAAGCAAAAGACAAATTCAGTCCCTTCTCCATG 480  
QY 481 AANTGTGACACGAGCAAACTGTTGGGAGAGACATTCGCAACAGAGATGGAACCCACAG 540  
Db 481 AANTGTGACACGAGCAAACTGTTGGGAGAGACATTCGCAACAGAGATGGAACCCACAG 540  
QY 541 GAGAACCAACAGAGAGGATGATGAGTCTTATGCGCATGATGATGATGATGATGATG 600  
Db 541 GAGAACCAACAGAGAGGATGATGAGTCTTATGCGCATGATGATGATGATGATGATG 600  
QY 601 AGACCCCTGGAACCTGAAGTATCTCATGAAGAAACCGAGCATAGTTACCACGTGGAAGA 660  
Db 601 AGACCCCTGGAACCTGAAGTATCTCATGAAGAAACCGAGCATAGTTACCACGTGGAAGA 660  
QY 661 CAGTTTCAAGAAGTATGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 720  
Db 661 CAGTTTCAAGAAGTATGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 720  
QY 721 ATTCCAGTGAACCAAGTAGTAGAAGATGCAACCATGATGATGATGATGATGATGATG 780  
Db 721 ATTCCAGTGAACCAAGTAGTAGAAGATGCAACCATGATGATGATGATGATGATGATG 780  
QY 781 ACCAAGTCTATGAGGAAACAGCAGATATAGAACTCTAGAAAAATGAAGGATGAATCA 840  
Db 781 ACCAAGTCTATGAGGAAACAGCAGATATAGAACTCTAGAAAAATGAAGGATGAATCA 840  
QY 841 CAGAGTAAGTCTCCCTGAGGATTAATCTCTGAGAGATTCACAGGTAATTTGTAAG 900  
Db 841 CAGAGTAAGTCTCCCTGAGGATTAATCTCTGAGAGATTCACAGGTAATTTGTAAG 900  
QY 901 AAGTAAGCATTTTCTGTGGAAGAACAGCAGGAAATGACCCAGAAACAAATAGAAAA 960  
Db 901 AAGTAAGCATTTTCTGTGGAAGAACAGCAGGAAATGACCCAGAAACAAATAGAAAA 960  
QY 961 CAGATGATCCAGAACAAAAAGTTAAGAAAAAGAGCCCTAACTTTTAAATTAAT 1020  
Db 961 CAGATGATCCAGAACAAAAAGTTAAGAAAAAGAGCCCTAACTTTTAAATTAAT 1020  
QY 1021 TTGATAAGACTATTAAGCTGAATGATGCTGCAAGAAACCTCCGTAAAGGGGAAAAA 1080  
Db 1021 TTGATAAGACTATTAAGCTGAATGATGCTGCAAGAAACCTCCGTAAAGGGGAAAAA 1080  
QY 1081 TTGAGGAGCAGTGAATGCAATTTAAAGAACTAGTACGCAAAATACCTTCAGAGTCCAG 1140  
Db 1081 TTGAGGAGCAGTGAATGCAATTTAAAGAACTAGTACGCAAAATACCTTCAGAGTCCAG 1140  
QY 1141 CAAGATATGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGTAATGAGG 1200  
Db 1141 CAAGATATGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGTAATGAGG 1200  
QY 1201 TGCTACCTGGAGCCATCGAGACCTTACCAAGAGGTGGCCAGCCTACCTGATGTCCTCGAG 1260  
Db 1201 TGCTACCTGGAGCCATCGAGACCTTACCAAGAGGTGGCCAGCCTACCTGATGTCCTCGAG 1260  
QY 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTCGTCAGACAGGCAACAATTTCTAGGTCAATGA 1320  
Db 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTCGTCAGACAGGCAACAATTTCTAGGTCAATGA 1320  
QY 1321 GAGTTTCCCTGCTTACCTTCAGAGATAGTTCAACTTTTCCCAATGATACTTCCCTAA 1380  
Db 1321 GAGTTTCCCTGCTTACCTTCAGAGATAGTTCAACTTTTCCCAATGATACTTCCCTAA 1380  
QY 1381 AAAATGACCTTGGCGTGGGATACCTCTTGTAGGAGATTAATGCAATGCAAGAAAGTTT 1440

Db 1381 AAAATGACCTTGGCTGGGATACCTCTTATAGAGATAATACAAATGCAAGAAAGTTT 1440  
Qy 1441 ATGAAGAGTGTGAGTGTGACACCTAATGATGGCTTTGCTAAATGCTCATTTGGTTCA 1500  
Db 1441 ATGAAGAGTGTGAGTGTGACACCTAATGATGGCTTTGCTAAATGCTCATTTGGTTCA 1500  
Qy 1501 TCTGAAGCCACAGAAACAAATTTGCTGAGAGCATCCATATTTAAAGGAAGGATAGAT 1560  
Db 1501 TCTGAAGCCACAGAAACAAATTTGCTGAGAGCATCCATATTTAAAGGAAGGATAGAT 1560  
Qy 1561 CCGGAGATCTGGCACTGATGATGGAGATTTTATTTCCACTCTGGGGATGCCATCAGA 1620  
Db 1561 CCGGAGATCTGGCACTGATGATGGAGATTTTATTTCCACTCTGGGGATGCCATCAGA 1620  
Qy 1621 GGGTTGGGAAACAAAGAGGCATATAAGTGTATGAGCTTTGGGCACAAAGAGAGACATTTG 1680  
Db 1621 GGGTTGGGAAACAAAGAGGCATATAAGTGTATGAGCTTTGGGCACAAAGAGAGACATTTG 1680  
Qy 1681 CATCTGTCTGGCAACGCTCACTCTCAATGTGAATGAGTCTTAAAGAGTCTTTCTGCTGAGGATG 1740  
Db 1681 CATCTGTCTGGCAACGCTCACTCTCAATGTGAATGAGTCTTAAAGAGTCTTTCTGCTGAGGATG 1740  
Qy 1741 CCCCAGAGAAACGGGCTCACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
Db 1741 CCCCAGAGAAACGGGCTCACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
Qy 1801 TCCGAGATGAAGGCTTTGCAAGTGTGATGGAATAAGCCAAAGGTCCTTCTGCTGAGGATG 1860  
Db 1801 TCCGAGATGAAGGCTTTGCAAGTGTGATGGAATAAGCCAAAGGTCCTTCTGCTGAGGATG 1860  
Qy 1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTCAAGCTGTGGCAGCAAGGAAGAA 1920  
Db 1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTCAAGCTGTGGCAGCAAGGAAGAA 1920  
Qy 1921 ATGAAATGCTGCAAGAGAGTCTTAAACCTGTACTTACAGAAAGTTTCCCGGAGA 1980  
Db 1921 ATGAAATGCTGCAAGAGAGTCTTAAACCTGTACTTACAGAAAGTTTCCCGGAGA 1980  
Qy 1981 CAACAGATGCAAGAGAGAGAGATCAATATTCATGATGCAACCCCGGAGATCAAGTGT 2040  
Db 1981 CAACAGATGCAAGAGAGAGAGATCAATATTCATGATGCAACCCCGGAGTCAAGTGT 2040  
Qy 2041 GGGCGCACAGGGCCCAAACTGCAAGGCTCCGAATGCACTGGGCTTGGTGAATCCCA 2100  
Db 2041 GGGCGCACAGGGCCCAAACTGCAAGGCTCCGAATGCACTGGGCTTGGTGAATCCCA 2100  
Qy 2101 AGGAAGGCTGCAAGATTCGATGTGCCAACGAGACAGGACCTGGGAGGAGGCAAGTGC 2160  
Db 2101 AGGAAGGCTGCAAGATTCGATGTGCCAACGAGACAGGACCTGGGAGGAGGCAAGTGC 2160  
Qy 2161 TCATCTTTGATGACTCTTTGAGCAGAGGATGGGAGATGCTCATCTTTCGGCTGA 2220  
Db 2161 TCATCTTTGATGACTCTTTGAGCAGAGGATGGGAGATGCTCATCTTTCGGGCTGA 2220  
Qy 2221 TATTCATGCTGATGCTGCTTGAAGTGTGGAATCGGAATCGACACAGAGAGAGCCCTCCAG 2280  
Db 2221 TATTCATGCTGATGCTGCTTGAAGTGTGGAATCGGAATCGACACAGAGAGAGCCCTCCAG 2280  
Qy 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTCTTGGAGAGA 2324  
Db 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTCTTGGAGAGA 2324

RESULT 4

US-09-903-063-3  
; Sequence 3, Application US/09903063  
; Patent No. 6815415  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.

; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV3  
; CURRENT APPLICATION NUMBER: US/09/903,063  
; CURRENT FILING DATE: 2001-10-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-903-063-3

Query Match 100.0%; Score 2324; DB 4; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CGGACCTGCAATGGCCAGCGCTAAGATGCGGAGAGAGCGGCAACAGCAGCAGCG 60  
Db 1 CGGACCTGCAATGGCCAGCGCTAAGATGCGGAGAGAGCGGCAACAGCAGCAGCG 60  
Qy 61 GCTCCGAGCGGTAGCAGCGAGTGGGCGAGCAGCAGCCCCGGGCGCCGAGAGAGACA 120  
Db 61 GCTCCGAGCGGTAGCAGCGAGTGGGCGAGCAGCAGCCCCGGGCGCCGAGAGAGACA 120  
Qy 121 AGCATGGAGGACACAAGAAATGGAGGAAAGGCGACTCTCGGGAAGCTTCATCTTCAGT 180  
Db 121 AGCATGGAGGACACAAGAAATGGAGGAAAGGCGACTCTCGGGAAGCTTCATCTTCAGT 180  
Qy 181 GGTATTGCTGATTCGATGCTGGCGCTCTGGACATCTGTAGCTGTGCTGCTTGGTTGATC 240  
Db 181 GGTATTGCTGATTCGATGCTGGCGCTCTGGACATCTGTAGCTGTGCTGCTTGGTTGATC 240  
Qy 241 TTGTTGATGATGAGGAAGTTCTAGGAAACCTAGGAACTATGATGCTGATGCTGATGAG 300  
Db 241 TTGTTGATGATGAGGAAGTTCTAGGAAACCTAGGAACTATGATGCTGATGCTGATGAG 300  
Qy 301 ATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
Db 301 ATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
Qy 361 CAGCAGTCCGCGCAGAGAGCTCAGCCACACATGCTGAGCCGAGAGAGAGTTCCTGTGG 420  
Db 361 CAGCAGTCCGCGCAGAGAGCTCAGCCACACATGCTGAGCCGAGAGAGAGTTCCTGTGG 420  
Qy 421 AGCAGAACCCAGAAATTCGAGATGAGAGAAACGAGATGATGATGATGATGATGATGATG 480  
Db 421 AGCAGAACCCAGAAATTCGAGATGAGAGAAACGAGATGATGATGATGATGATGATGATG 480  
Qy 481 AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540  
Db 481 AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540  
Qy 541 GAGAACCAACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
Db 541 GAGAACCAACAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
Qy 601 AGACCTGGACCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660  
Db 601 AGACCTGGACCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660  
Qy 661 CAGTTTCAAGAGCTGTAATCAGGATATGGAAGAGATGATGATGATGATGATGATGATGATG 720  
Db 661 CAGTTTCAAGAGCTGTAATCAGGATATGGAAGAGATGATGATGATGATGATGATGATGATG 720  
Qy 721 ATTCCAGTGAACAGAGTATGAGAAAGATGCAACATGATGATGATGATGATGATGATGATG 780  
Db 721 ATTCCAGTGAACAGAGTATGAGAAAGATGCAACATGATGATGATGATGATGATGATGATG 780  
Qy 781 ACCAAGTCTATGAGGAAACAGAGTATGATGATGATGATGATGATGATGATGATGATGATG 840  
Db 781 ACCAAGTCTATGAGGAAACAGAGTATGATGATGATGATGATGATGATGATGATGATGATG 840

QY	841	CAGAAGTAAC	TGCTCCCTGAGGATAA	TCTGTAGAAAGAT	TCACAGTAAT	TGTGAAG	900
DB	841	CAGNAGTAAC	TGCTCCCTGAGGATAA	TCTGTAGAAAGAT	TCACAGTAAT	TGTGAAG	900
QY	901	AAGTAAGCA	TTTTTCTGTGGAAGAAC	CACAGCAAGT	ACCACGAAACA	AAATAGAAAA	960
DB	901	AAGTAAGCA	TTTTTCTGTGGAAGAAC	CACAGCAAGT	ACCACGAAACA	AAATAGAAAA	960
QY	961	CAGATGAT	TCAGAAACAAAAGC	AAAGTTAA	AAAAAAGAAAGC	CTTAAATAAAT	1020
DB	961	CAGATGAT	TCAGAAACAAAAGC	AAAGTTAA	AAAAAAGAAAGC	CTTAAATAAAT	1020
QY	1021	TTGATAAGCA	CTATTAAGCTGAAC	CTTGATGCTGCGAAAAAA	AACTCCGTAAAAAGGGAAAAA	1080	
DB	1021	TTGATAAGCA	CTATTAAGCTGAAC	CTTGATGCTGCGAAAAAA	AACTCCGTAAAAAGGGAAAAA	1080	
QY	1081	TTGAGGNAAG	CAGTGAATGCAATTTAA	AGAACTAGTACGCAAA	TACCCTCAGAGTCCACGAG	1140	
DB	1081	TTGAGGNAAG	CAGTGAATGCAATTTAA	AGAACTAGTACGCAAA	TACCCTCAGAGTCCACGAG	1140	
QY	1141	CAAGATAT	GGAAAGCGCAGTGTGAGGAT	GATTTGGCTGAGAAGAGGAG	AAGTAATGAGG	1200	
DB	1141	CAAGATAT	GGAAAGCGCAGTGTGAGGAT	GATTTGGCTGAGAAGAGGAG	AAGTAATGAGG	1200	
QY	1201	TGCTACGTGAG	CCATCGACACCTTAC	CAAGAGGTGCGCAGCCT	ACTGATGTCCTGCGAG	1260	
DB	1201	TGCTACGTGAG	CCATCGACACCTTAC	CAAGAGGTGCGCAGCCT	ACTGATGTCCTGCGAG	1260	
QY	1261	ACCTGCTGAAG	CTGAGTTTGAAGCGTGGCT	CAGACAGGCAACA	ATTTCTGAGTCAATGA	1320	
DB	1261	ACCTGCTGAAG	CTGAGTTTGAAGCGTGGCT	CAGACAGGCAACA	ATTTCTGAGTCAATGA	1320	
QY	1321	GAGGTTCCCTG	CTTACCCTCGAGAGAT	TAGTTCAA	CTATTTCCCAGTATCTTCCCTTAA	1380	
DB	1321	GAGGTTCCCTG	CTTACCCTCGAGAGAT	TAGTTCAA	CTATTTCCCAGTATCTTCCCTTAA	1380	
QY	1381	AAAATGAC	TTGGCTGGGATAC	CTCTTGATAGGAGAT	AATGACAAATGCAAGAAAGTTT	1440	
DB	1381	AAAATGAC	TTGGCTGGGATAC	CTCTTGATAGGAGAT	AATGACAAATGCAAGAAAGTTT	1440	
QY	1441	ATGAAGAGG	TGCTGAGTGTGAC	ACCTAATGATGGCTT	TGCTAAAGTCCATATGCGCTTCA	1500	
DB	1441	ATGAAGAGG	TGCTGAGTGTGAC	ACCTAATGATGGCTT	TGCTAAAGTCCATATGCGCTTCA	1500	
QY	1501	TCCTGAAGGC	CAGACAAATTCCTGAGAGCA	TCCCATATTTAAAGGAAGGA	ATAAGAT	1560	
DB	1501	TCCTGAAGGC	CAGACAAATTCCTGAGAGCA	TCCCATATTTAAAGGAAGGA	ATAAGAT	1560	
QY	1561	CCGGAGAT	CTCTGGCACTGATCGGAGAT	TTTTATTTCC	ACCTGGGGATGCCATGCAGA	1620	
DB	1561	CCGGAGAT	CTCTGGCACTGATCGGAGAT	TTTTATTTCC	ACCTGGGGATGCCATGCAGA	1620	
QY	1621	GGGTTGGGA	CAAGAGGCATAT	TAAGTGTATGAGCTT	GGGCAACAAGAGGACAC	CTTTG	1680
DB	1621	GGGTTGGGA	CAAGAGGCATAT	TAAGTGTATGAGCTT	GGGCAACAAGAGGACAC	CTTTG	1680
QY	1681	CATCTGTCTGG	CAACGCTCACTCTAC	AAATGTGAATGGA	CTCAAGCACACGCTTGTGGGA	1740	
DB	1681	CATCTGTCTGG	CAACGCTCACTCTAC	AAATGTGAATGGA	CTCAAGCACACGCTTGTGGGA	1740	
QY	1741	CCCCAAAAA	ACGGGCTACACAGAGTT	AGTAAAGTCTTTT	AGAAAGAACTGGAGTTAA	1800	
DB	1741	CCCCAAAAA	ACGGGCTACACAGAGTT	AGTAAAGTCTTTT	AGAAAGAACTGGAGTTAA	1800	
QY	1801	TCCGAGAT	CAAGGCCTTGCAGTGA	TGATAAAGCCAAAGG	TCTCTTCCCTGCTGAGGATG	1860	
DB	1801	TCCGAGAT	CAAGGCCTTGCAGTGA	TGATAAAGCCAAAGG	TCTCTTCCCTGCTGAGGATG	1860	
QY	1861	AAAACTGAGG	AAAAAGGGGACTGGAG	CGAGTTAC	CGCTGTGGCAGCAAGGAGAGAA	1920	
DB	1861	AAAACTGAGG	AAAAAGGGGACTGGAG	CGAGTTAC	CGCTGTGGCAGCAAGGAGAGAA	1920	

QY	1921	ATGAAATGCTGCAAGAGGAGCTCTCTAAACCTGTACCTTACTAGAAAAAGTTCCCGGAGA	1980
Db	1921	ATGAAATGCTGCAAGAGGAGCTCTCTAAACCTGTACCTTACTAGAAAAAGTTCCCGGAGA	1980
QY	1981	CAACAGAGTCGAAAGAGGAGCAGATCAATATATTCATCATGACACCCCGGAGCTCACGTGT	2040
Db	1981	CAACAGAGTCGAAAGAGGAGCAGATCAATATATTCATCATGACACCCCGGAGCTCACGTGT	2040
QY	2041	GGCGGCACACAGGGCCACAAAATCGCAGGCTCCGAATTCACCTTGGCTTGTGTATCCCA	2100
Db	2041	GGCGGCACACAGGGCCACAAAATCGCAGGCTCCGAATTCACCTTGGCTTGTGTATCCCA	2100
QY	2101	AGGAAGGCTCGAAGATTCGATGTGCCAACAGACACAGGACCTGGGAGGAGGCAAGGTGC	2160
Db	2101	AGGAAGGCTCGAAGATTCGATGTGCCAACAGACACAGGACCTGGGAGGAGGCAAGGTGC	2160
QY	2161	TCATCTTTTGATGACTCCTTTTGACACAGAGGTATGGCAGGATGCTCATCTTTCCGGCTGA	2220
Db	2161	TCATCTTTTGATGACTCCTTTTGACACAGAGGTATGGCAGGATGCTCATCTTTCCGGCTGA	2220
QY	2221	TATTCATCTGGATGTGTGGCATCCGGAACTGACACACAGCAGACGCGCTTCCAG	2280
Db	2221	TATTCATCTGGATGTGTGGCATCCGGAACTGACACACAGCAGACGCGCTTCCAG	2280
QY	2281	CAATTTAGCATGAATTCATGCAAGCTTTGGGAACTCTGGAGAGA	2324
Db	2281	CAATTTAGCATGAATTCATGCAAGCTTTGGGAACTCTGGAGAGA	2324

RESULT 5

US-09-859-604-3

; Sequence 3, Application US/09859604

; Patent No. 6835370

; GENERAL INFORMATION:

; APPLICANT: Wands, Jack R.

; APPLICANT: de la Monte, Suzanne M

; APPLICANT: Deutch, Allan H

; APPLICANT: Ghanbari, Hosein A

; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS

; FILE REFERENCE: 21486-032 CIP

; CURRENT APPLICATION NUMBER: US/09/859,604

; PRIOR FILING DATE: 2001-05-17

; PRIOR APPLICATION NUMBER: 09/436,184

; PRIOR FILING DATE: 1999-11-08

; NUMBER OF SEQ ID NOS: 13

; SOFTWARE: PatentIn Ver.. 2.1

; SEQ ID NO: 3

; LENGTH: 2324

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-859-604-3

	Query Match	100.0%;	Score 2324;	DB 4;	Length 2324;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 2324;	Conservative	0;	Mismatches	0;
				Indels	Gaps
				0;	0;
QY	1	CGGACCGTGCATGGCCCGCAGCGTAAAGTCCCAAGCAGCGCGGCAACAGCAGCAGCGG	60		
Db	1	CGGACCGTGCATGGCCCGCAGCGTAAAGTCCCAAGCAGCGCGGCAACAGCAGCAGCGG	60		
QY	61	GCTCCGGCAGCGTAGCAGCGTGGCGGCGACAGCAGCGCCCGGGCCCGGAGAGAGACAA	120		
Db	61	GCTCCGGCAGCGTAGCAGCGTGGCGGCGACAGCAGCGCCCGGGCCCGGAGAGAGACAA	120		
QY	121	AGCATGGAGGACACAAAGATGGAGGAAAGCGGAGCTCTCGGGAATTCATCTTCACGT	180		
Db	121	AGCATGGAGGACACAAAGATGGAGGAAAGCGGAGCTCTCGGGAATTCATCTTCACGT	180		
QY	181	GGTTTATGGTGAATTCATCTGGGGCTCTGGACATCTCTAGCTGTCTGTTTGGTTGATC	240		
Db	181	GGTTTATGGTGAATTCATCTGGGGCTCTGGACATCTCTAGCTGTCTGTTTGGTTGATC	240		
QY	241	TTGTTGATCATAGGAAAGTTCTTAGAAAAACCTAGGAATCTATGATGCTGATGGTGGAG	300		







```

1  TITLE OF INVENTION:  A GENE ENCODING A NOVEL MARKER FOR
2  TITLE OF INVENTION:  CANCER
3  NUMBER OF SEQUENCES:  9
4  CORRESPONDENCE ADDRESS:
5  ADDRESSEE:  BRINKS, HOFER, GILSON & LIONE
6  STREET:  NBC Tower - Suite 3600, 455 N. Cityfront
7  STREET:  Plaza Drive
8  CITY:  Chicago
9  STATE:  Illinois
10 COUNTRY:  USA
11 ZIP:  60611-5599
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE:  Floppy disk
14 COMPUTER:  IBM PC compatible
15 OPERATING SYSTEM:  PC-DOS/MS-DOS
16 SOFTWARE:  Patent In Release #1.0, Version #1.30
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER:  US/09/040,485
19 FILING DATE:  17-MAR-1998
20 CLASSIFICATION:  530
21 ATTORNEY/AGENT INFORMATION:
22 NAME:  Martin, Alice O.
23 REGISTRATION NUMBER:  35,601
24 REFERENCE/DOCKET NUMBER:  8998/3
25 TELECOMMUNICATION INFORMATION:
26 TELEPHONE:  312-321-4200
27 TELEFAX:  312-321-4299
28 INFORMATION FOR SEQ ID NO:  1:
29 SEQUENCE CHARACTERISTICS:
30 LENGTH:  2442 base pairs
31 TYPE:  nucleic acid
32 STRANDEDNESS:  single
33 TOPOLOGY:  linear
34 MOLECULE TYPE:  cDNA
35 FEATURE:
36 NAME/KEY:  CDS
37 LOCATION:  70..834
38 US-09-040-485-1

```

Query Match	35.4%;	Score	823.8;	DB	3;	Length	2442;
Best Local Similarity	99.2%;	Pred. No.	6.8e-226;				
Matches	828;	Conservative	0;	Mismatches	7;	Indels	0;
Gaps	0;						

  

QY	121	AGCATGGAGGACACAAGAATGGGAGGAAAGCGGCACTCTCGGGAACCTTCATCTTCACGT	180
Db	5	AGCTTGAAGGACACAAGAATGGGAGGAAAGCGGCACTCTCAGGAACCTCATCTTCACGT	64
QY	181	GGTTTATGGTGATTGCATTGCTGGCGCTGCGACATCTGTAGCTGTGCTTTGGTTTGATC	240
Db	65	GGTTTATGGTGATTGCATTGCTGGCGCTGCGACATCTGTAGCTGTGCTTTGGTTTGATC	124
QY	241	TTGTTGACTATGAGGAAGTCTTAGGAAACTTAGGAATCTATGATGCTGATGGTGTAGGAG	300
Db	125	TTGTTGACTATGAGGAAGTCTTAGGAAACTTAGGAATCTATGATGCTGATGGTGTAGGAG	184
QY	301	ATTTTGTATGGTATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTTCAGAGC	360
Db	185	ATTTTGTATGGTATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTTCAGAGC	244
QY	361	CAGCAGTCCCGCAGGAAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCCTGTGG	420
Db	245	CAGCAGTCCCGCAGGAAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCCTGTGG	304
QY	421	AGGCAGAACCCCAGAAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG	480
Db	305	AGGCAGAACCCCAGAAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG	364
QY	481	AAATGGTACACCGAACAATGTTGAGGNGAGACCTGCCAACAGAAGATGCCACACAG	540
Db	365	AAATGGTACACCGAACAATGTTGAGGNGAGACCTTGCACAGAGATGACCCACAG	424
QY	541	GAGAAACCAACAAGAGGATGATGAGTTCTTTATGGCGACTGATGATGATGATGATG	600

```

RESULT 7
US-09-702-705-1791
; Sequence 1791, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Lique
; TITLE OF INVENTION: COMPOSITIONS AND
; TITLE OF INVENTION: DIAGNOSIS OF LUN
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Versio
; SEQ ID NO 1791
; LENGTH: 2442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-702-705-1791

```

Query Match	35.2%	Score	819	DB	4	Length	2442
Best Local Similarity	98.8%	Pred.	No. 1.6e-234				
Matches	825	Conservative	0	Mismatches	10	Indels	0
Gaps	0						
QY	121	AGCATGGAGGACACAGAATCGGAGGAGAAAGCGGACTCTCGGGAACCTTCATTTCTCAAGT	180				
Db	5	AGCTTGAGGACACAGAAATCGGAGGAGAAAGCGGACTCTCAGGAACCTTCATTTCTCAAGT	64				
QY	181	GGTTTATGGTGATTCGATTCGTGGGCGTCTGGACATCTGTAGCTGTGCTTTGGTTTGATC	240				
Db	65	GGTTTATGGTGATTCGATTCGTGGGCGTCTGGACATCTGTAGCTGTGCTTTGGTTTGATC	124				
QY	241	TTGTTGACTATCAGGAAGCTCTAGGAAAACCTAGGAATCTATGATGCTGATGGTGATGGAG	300				
Db	125	TTGTTGACTATCAGGAAGTCTTAGGAAAACCTAGGAATCTATGATGCTGATGGTGATGGAG	184				
QY	301	ATTTTGATGTGATGATGCCAAAGTTTTATTAGGACTTAAGAGAGATCTACTTTCAGAGC	360				
Db	185	ATTTTGATGTGATGATGCCAAAGTTTTATTAGGACTTAAGAGAGATCTACTTTCAGAGC	244				

Qy 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTGTGG 420  
Db 245 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTGTGG 304  
Qy 421 AGCGAGAACCCCGAGAAATATCGAAGATGAAGCAAAAGCAAAATTCAGTCCCTCCATG 480  
Db 305 AGCGAGAACCCCGAGAAATATCGAAGATGAAGCAAAAGCAAAATTCAGTCCCTCCATG 364  
Qy 481 AAATGGTACGCGAGAACATGTTGAGGAGAGACTTGCACAGAGAGATGGACCCACAG 540  
Db 365 AAATGGTACGCGAGAACATGTTGAGGAGAGACTTGCACAGAGAGATGGACCCACAG 424  
Qy 541 GAGAACCAACAGAGAGATGATGAGTTCTTATGGCGACTGATGATGATGATGATG 600  
Db 425 GAGAACCAACAGAGAGATGATGAGTTCTTATGGCGACTGATGATGATGATGATG 484  
Qy 601 AGACCTGGAACTCGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTGGAAGAGA 660  
Db 485 AGACCTGGAACTCGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTGGAAGAGA 544  
Qy 661 CAGTTTCAAGAGCTGTAATCAGGATATGAGAGATGATGATGATGATGATGATG 720  
Db 545 CAGTTTCAAGAGCTGTAATCAGGATATGAGAGATGATGATGATGATGATGATG 604  
Qy 721 ATTCCAGTGAACGAGTATGAGAGATGAGAAAGTTGCAACCATGATACAGATGATG 780  
Db 605 ATTCCAGTGAACGAGTATGAGAGATGAGAAAGTTGCAACCATGATACAGATGATG 664  
Qy 781 ACCAAGTCTATGAGAAACAGCAGTATGTAACCTCTAGAAAATGAAGGATGAATCA 840  
Db 665 ACCAAGTCTATGAGAAACAGCAGTATGTAACCTCTAGAAAATGAAGGATGAATCA 724  
Qy 841 CAGAAGTAACTGCTCCCTCGAGGATAATCTGTAGAGATTCACAGGTAAATGTAGAG 900  
Db 725 CAGAAGTAACTGCTCCCTCGAGGATAATCTGTAGAGATTCACAGGTAAATGTAGAG 784  
Qy 901 AAGTAAGCATTTTCTGTGGAAGAACAGCAGGAGTACCCAGGAAACAAATAG 955  
Db 785 AAGTAAGCATTTTCTGTGGAAGAACAGCAGGAGTACCCAGGAGTACCTTAAAG 839

## RESULT 8

US-09-736-457-1791  
; Sequence 1791, Application US/09736457  
; Patent No. 6509448  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Aijun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C15  
; CURRENT APPLICATION NUMBER: US/09/736,457  
; NUMBER OF SEQ ID NOS: 1864  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1791  
; LENGTH: 2442  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-736-457-1791

Query Match 35.2%; Score 819; DB 4; Length 2442;  
Best Local Similarity 98.8%; Pred. No. 1.6e-224;  
Matches 825; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 121 AGCATGGAGGACACAGAGATGGAGGAAAGGCGGACTCTCGGGAATTCATTCTTCACTG 180  
Db 5 AGCTTGAAGGACACAGAGATGGAGGAAAGGCGGACTCTCAGGAATTCATTCTTCACTG 64  
Qy 181 GGTTCATGGTGAATGCAATTCCTGGCGCTCTGGACATCTGTAGCTGTGTGTTGATC 240  
Db 65 GGTTCATGGTGAATGCAATTCCTGGCGCTCTGGACATCTGTAGCTGTGTGTTGATC 124  
Qy 241 TTGTGACTATGAGGAGGTTCTAGGAAATCTAGGAATCTATGATGCTGATGATGAG 300  
Db 125 TTGTGACTATGAGGAGGTTCTAGGAAATCTAGGAATCTATGATGCTGATGATGAG 184  
Qy 301 ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
Db 185 ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 244  
Qy 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTGTGG 420  
Db 245 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTGTGG 304  
Qy 421 AGCGAGAACCCCGAGAAATATCGAAGATGAGCAAAAGAAACAAATTCAGTCCCTTCCATG 480  
Db 305 AGCGAGAACCCCGAGAAATATCGAAGATGAGCAAAAGAAACAAATTCAGTCCCTTCCATG 364  
Qy 481 AAATGGTACACGCGAGAACATGTTGAGGAGAGAGACTTGCACAGAGAGATGGACCCACAG 540  
Db 365 AAATGGTACACGCGAGAACATGTTGAGGAGAGAGACTTGCACAGAGAGATGGACCCACAG 424  
Qy 541 GAGAACCAACAGAGAGATGATGAGTTCTTATGGCGACTGATGATGATGATGATGATG 600  
Db 425 GAGAACCAACAGAGAGATGATGAGTTCTTATGGCGACTGATGATGATGATGATGATG 484  
Qy 601 AGACCTGGAACTCGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTGGAAGAGA 660  
Db 485 AGACCTGGAACTCGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTGGAAGAGA 544  
Qy 661 CAGTTTCAAGAGCTGTAATCAGGATATGAGAGATGATGATGATGATGATGATG 720  
Db 545 CAGTTTCAAGAGCTGTAATCAGGATATGAGAGATGATGATGATGATGATGATG 604  
Qy 721 ATTCCAGTGAACGAGTATGAGAGATGAGAAAGTTGCAACCATGATACAGATGATG 780  
Db 605 ATTCCAGTGAACGAGTATGAGAGATGAGAAAGTTGCAACCATGATACAGATGATG 664  
Qy 781 ACCAAGTCTATGAGAAACAGCAGTATGTAACCTCTAGAAAATGAAGGATGAATCA 840  
Db 665 ACCAAGTCTATGAGAAACAGCAGTATGTAACCTCTAGAAAATGAAGGATGAATCA 724  
Qy 841 CAGAAGTAACTGCTCCCTCGAGGATAATCTGTAGAGATTCACAGGTAAATGTAGAG 900  
Db 725 CAGAAGTAACTGCTCCCTCGAGGATAATCTGTAGAGATTCACAGGTAAATGTAGAG 784  
Qy 901 AAGTAAGCATTTTCTGTGGAAGAACAGCAGGAGTACCCAGGAAACAAATAG 955  
Db 785 AAGTAAGCATTTTCTGTGGAAGAACAGCAGGAGTACCCAGGAGTACCTTAAAG 839

## RESULT 9

US-09-671-325-1791  
; Sequence 1791, Application US/09671325  
; Patent No. 6667154  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1791
; LENGTH: 2442
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-671-325-1791

Query Match 35.2%; Score 819; DB 4; Length 2442;
Best Local Similarity 98.8%; Pred. No. 1.6e-224; Indels 0; Gaps 0;
Matches 825; Conservative 0; Mismatches 10;

QY 121 AGCATGGAGGACACAAAGAAATGGAGGAAAGCGGACTCTCGGAACTTCACTTTCACGT 180
DB 5 AGCTTGAAGGACACAAAGAAATGGAGGAAAGCGGACTCTCAGGAACCTTCACTTTCACGT 64

QY 181 GGTATATGGTGAATGCTGCGGCGTCTGGACATCTGTAGCTGTCTGTTTGGTTGATC 240
DB 65 GGTATATGGTGAATGCTGCGGCGTCTGGACATCTGTAGCTGTCTGTTTGGTTGATC 124

QY 241 TTGTTGACTATGAGGAAGTCTAGGAAACTAGGAATCTATGATGCTGATGGTGGAG 300
DB 185 ATTTTGAAGGACACAAAGAAATGGAGGAAAGCGGACTCTCAGGAACCTTCACTTTCACGT 244

QY 301 ATTTTGAAGGACACAAAGAAATGGAGGAAAGCGGACTCTCAGGAACCTTCACTTTCACGT 360
DB 185 ATTTTGAAGGACACAAAGAAATGGAGGAAAGCGGACTCTCAGGAACCTTCACTTTCACGT 244

QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGGAGTCTCTCTGG 420
DB 245 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGGAGTCTCTCTGG 304

QY 421 AGCCAGAACCCCAAGAGATGATGATGCTTCTTATGGCGACTGATGATGATGATGATG 480
DB 305 AGCCAGAACCCCAAGAGATGATGATGCTTCTTATGGCGACTGATGATGATGATGATG 364

QY 481 AAATGGTACACGAGAACTGTTGAGGAGAAAGTCTGCAACAAAGAGATGGAGCCACAG 540
DB 365 AAATGGTACACGAGAACTGTTGAGGAGAAAGTCTGCAACAAAGAGATGGAGCCACAG 424

QY 541 GAGAACCAACAAAGAGGATGATGATGCTTCTTATGGCGACTGATGATGATGATGATG 600
DB 425 GAGAACCAACAAAGAGGATGATGATGCTTCTTATGGCGACTGATGATGATGATGATG 484

QY 601 AGACCCCTGGAACTTGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTTGGAAGA 660
DB 485 AGACCCCTGGAACTTGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTTGGAAGA 544

QY 661 CAGTTTCAACAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
DB 545 CAGTTTCAACAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 604

QY 721 ATTCCAGTGAACCAAGTAGTAGAGATGAAAGATGACCATGATGATGATGATGATGATG 780
DB 605 ATTCCAGTGAACCAAGTAGTAGAGATGAAAGATGACCATGATGATGATGATGATGATG 664

QY 781 ACCAAGTCTATGAGGAACCAAGTAGTAGAGATGAAAGATGACCATGATGATGATGATGATG 840
DB 665 ACCAAGTCTATGAGGAACCAAGTAGTAGAGATGAAAGATGACCATGATGATGATGATGATG 724

QY 841 CAGAAGTAACTGCTCCCTGAGGATAATCTCTGTAAGATTCACAGGTAATTTGTAGAG 900
DB 725 CAGAAGTAACTGCTCCCTGAGGATAATCTCTGTAAGATTCACAGGTAATTTGTAGAG 784

QY 901 AAGTAAGCAATTTTCTCTGTGGAAAGACAGCAGGAAGTACCAACAGAAACCAATATG 955
DB 785 AAGTAAGCAATTTTCTCTGTGGAAAGACAGCAGGAAGTACCAACAGGAAGTACCAACAGGA 839

RESULT 10
US-09-659-521-1
; Sequence 1, Application US/09659521
; Patent No. 6727080
; GENERAL INFORMATION:
; APPLICANT: RADOSEVICH, JAMES A.
; TITLE OF INVENTION: GENE ENCODING LABYRINTHIN, A MARKER FOR CANCER
; FILE REFERENCE: 21511/91057
; CURRENT APPLICATION NUMBER: US/09/659,521
; CURRENT FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: PCT/US99/05365
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: 09/040,485
; PRIOR FILING DATE: 1998-03-17
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2442
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (70)...(834)
US-09-659-521-1

Query Match 35.2%; Score 819; DB 4; Length 2442;
Best Local Similarity 98.8%; Pred. No. 1.6e-224; Indels 0; Gaps 0;
Matches 825; Conservative 0; Mismatches 10;

QY 121 AGCATGGAGGACACAAAGAAATGGAGGAAAGCGGACTCTCGGAACTTCACTTTCACGT 180
DB 5 AGCTTGAAGGACACAAAGAAATGGAGGAAAGCGGACTCTCAGGAACCTTCACTTTCACGT 64

QY 181 GGTATATGGTGAATGCTGCGGCGTCTGGACATCTGTAGCTGTCTGTTTGGTTGATC 240
DB 65 GGTATATGGTGAATGCTGCGGCGTCTGGACATCTGTAGCTGTCTGTTTGGTTGATC 124

QY 241 TTGTTGACTATGAGGAAGTCTAGGAAACTAGGAATCTATGATGCTGATGGTGGAG 300
DB 125 TTGTTGACTATGAGGAAGTCTAGGAAACTAGGAATCTATGATGCTGATGGTGGAG 184

QY 301 ATTTTGAAGGACACAAAGAAATGGAGGAAAGCGGACTCTCAGGAACCTTCACTTTCACGT 360
DB 185 ATTTTGAAGGACACAAAGAAATGGAGGAAAGCGGACTCTCAGGAACCTTCACTTTCACGT 244

QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGGAGTCTCTCTGG 420
DB 245 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGGAGTCTCTCTGG 304

QY 421 AGCCAGAACCCCAAGAGATGATGATGCTTCTTATGGCGACTGATGATGATGATGATG 480
DB 305 AGCCAGAACCCCAAGAGATGATGATGCTTCTTATGGCGACTGATGATGATGATGATG 364

QY 481 AAATGGTACACGAGAACTGTTGAGGAGAAAGTCTGCAACAAAGAGATGGAGCCACAG 540
DB 365 AAATGGTACACGAGAACTGTTGAGGAGAAAGTCTGCAACAAAGAGATGGAGCCACAG 424

QY 541 GAGAACCAACAAAGAGGATGATGATGCTTCTTATGGCGACTGATGATGATGATGATG 600
DB 425 GAGAACCAACAAAGAGGATGATGATGCTTCTTATGGCGACTGATGATGATGATGATG 484

QY 601 AGACCCCTGGAACTTGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTTGGAAGA 660
DB 485 AGACCCCTGGAACTTGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTTGGAAGA 544

QY 661 CAGTTTCAACAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
DB 545 CAGTTTCAACAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 604

QY 721 ATTCCAGTGAACCAAGTAGTAGAGATGAAAGATGACCATGATGATGATGATGATGATG 780
DB 605 ATTCCAGTGAACCAAGTAGTAGAGATGAAAGATGACCATGATGATGATGATGATGATG 664

QY 781 ACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAATGAAGGATGAATCA 840  
Db 665 ACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAATGAAGGATGAATCA 724  
QY 841 CAGAAGTAACTGCTCCCTGAGGATAATCTCTAGAGATTCAAGGATTAATGTAAG 900  
Db 725 CAGAAGTAACTGCTCCCTGAGGATAATCTCTAGAGATTCAAGGATTAATGTAAG 784  
QY 901 AAGTAAGCATTTTCTCTGGAAGAACAGCAGAGTACCAAGGATTAATGTAAG 955  
Db 785 AAGTAAGCATTTTCTCTGGAAGAACAGCAGAGTACCAAGGATTAATGTAAG 839

RESULT 11  
US-09-949-016-3447/c  
; Sequence 3447, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3447  
; LENGTH: 956  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-3447

Query Match 23.2%; Score 539.8; DB 4; Length 956;  
Best Local Similarity 83.3%; Pred. No. 1.7e-144;  
Matches 782; Conservative 0; Mismatches 37; Indels 120; Gaps 10;

QY 111 ACAGAGCAAAAGCATGGAGGACACAGAAATGGAGGAAAGCGGACTCTCGGAACTTCA 170  
Db 881 ACAGAGCAAAAGCATGGAGGACACAGAAATGGAGGAAAGCGGACTCTCGGAACTTCA 822  
QY 171 TTCTTCAGTGGTTTATGGTGAATGCAATGCTGGCGCTCTGGACATCTGTAGTGTGTT 230  
Db 821 TTCTTCAGTGGTTTATGGTGAATGCAATGCTGGCGCTCTGGACATCTGTAGTGTGTT 762  
QY 231 TGGTTTCATCTCTGTCATCTATGAGGAAGTTCTAGGAAATCTAGGAATCTATGATGAT 290  
Db 761 TGGTTTCATCTCTGTCATCTATGAGGAAGTTCTAGGAAATCTAGGAATCTATGATGAT 712  
QY 291 GGTGATCGGATTTTTCATCTGATGATGATGATGATGATGATGATGATGATGATGAT 331  
Db 711 GGTGATCGGATTTTTCATCTGATGATGATGATGATGATGATGATGATGATGATGAT 652  
QY 332 ---AGGACTTAAAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTTGAGCC 388  
Db 651 TGCAGGATTTTAAAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTTGAGCC 592  
QY 389 ACACATGAGCCCGAGGAGCAGGTTCTCTGTGGAGGAG--AACCCGAGAAATATGAGAT 446  
Db 591 ACACATGAGCCCGAGGAGCAGGTTCTCTGTGGAGGAGGTAAGCCAGAAATATGAGAT 532  
QY 447 GAAGCAAAAGCAAAATTCAGTCCCTCTCCATGAATGTGTACGAGCAGCAGATGTTGAG 506  
Db 531 GAAGCAAAAGCAAAATTCAGTCCCTCTCCATGAATGTGTACGAGCAGCAGATGTTGAG 472  
QY 507 GGAGNAGATTCTGCAACAGAGATGGACCCACAGGAGAACCCACAAAGAGATGATGAG 566  
Db 471 AATTAAATTTGCAACAGAGATGGACCCACAGGAGAACCCACAAAGAGATGATGAG 412

QY 567 TTCTTATGGCGACTGATGTAGATGATGATGATGATGATGATGATGATGATGATGAT 626  
Db 411 TTCTTATGGCGACTGATGTAGATGATGATGATGATGATGATGATGATGATGATGAT 352  
QY 627 GA---AGAAACCGAGCATAGTTTACACGCTGGAAGAGACAG-----TTTCAAGAC 674  
Db 351 GAAGGTAGAACACAGCATAGTTTACACGCTGGAAGAGACAGGTGAGATTTTAAATTT 292  
QY 675 TGTATCAGGATATGGAAGAGATGATGCTGAGCAGGAGAAATCCAGATTCAGTGAACCA 734  
Db 291 TGTATCAGGATATGGAAGAGATGATGCTGAGCAGGAGAAATCCAGATTCAGTGAACCA 262  
QY 735 GTAGTAGAAGATGAAGATTGACCATGATACAGATGATGATACATACCAAGTCTATGAG 794  
Db 261 GTAGTAGAAGATGAAGATTGACCATGATACAG-----CAAGTCTATGAG 216  
QY 795 GAACAAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCAG-----AAGTAAC 850  
Db 215 GAACAAGTGTGCAATGCTCTAGAAATGAAGGATAGAAATCAGAGTAAGTTTAT 156  
QY 851 TGCTCCCTCTGAGGATAATCTCTGTAAGATTCACAGGTAATTTGTAGA-----898  
Db 155 TGCTCCCTCTGAGGATAATCTCTGTAAGATTCACAGGTAATTTGTAGA-----96  
QY 899 -----AGAAAGTAAGCATTTTCTCTGTAAGATTCACAGGTAATTTGTAGA-----948  
Db 95 TATTTATTTTGAAGTAAGCATTTTCTCTGTAAGATTCACAGGTAATTTGTAGA-----39  
QY 949 CAATAGAAAAACAGATGATCCAGAACCAAAAGCAGAAAG 987  
Db 38 -AATAGAAAAACAGATGATCCAGAACCAAAAGCAGAAAG 1

RESULT 12  
US-09-949-016-15189/c  
; Sequence 15189, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15189  
; LENGTH: 74730  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(74730)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-15189

Query Match 6.6%; Score 153.2; DB 4; Length 74730;  
Best Local Similarity 98.1%; Pred. No. 7.3e-32;  
Matches 155; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 111 AGAGAGCAAAAGCATGGAGGACACAGAAATGGAGGAAAGCGGACTCTCGGAACTTCA 170  
Db 67174 ACAGAGCAAAAGCATGGAGGACACAGAAATGGAGGAAAGCGGACTCTCAGGAATTC 67115  
QY 171 TTCTTACGTGTTTATGGTGAATGATGATGATGATGATGATGATGATGATGATGAT 230

Db 67114 TTCTTACGTTGGTTATGTTGATTCCTGGGCGTCTGGACATCTGTAGCTGCTGTT 67055

QY 231 TGGTTTCATCTTGTGTACTATGAGGAAGTTCTTAGGAAA 268

Db 67054 TGGTTTGATCTTGTGTACTATGAGGAAGTTCTAGGTAA 67017

RESULT 13

US-09-949-016-16453/c

Sequence 16453, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 16453

LENGTH: 5746

TYPE: DNA

ORGANISM: Human

US-09-949-016-16453

Query Match 3.8%; Score 88.8; DB 4; Length 5746;

Best Local Similarity 97.8%; Pred. No. 3.8e-14;

Mismatches 0; Indels 0; Gaps 0;

Matches 90; Conservative 0;

QY 1069 AAGGGGAAAATTCAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGCGAAATACCCCTC 1128

Db 700 AACAGGGAAAATTCAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGCGAAATACCCCTC 641

QY 1129 AGAGTCCACGACGACAGATATGGAAGGCGCAG 1160

Db 640 AGAGTCCACGACGACAGATATGGAAGGCGCAG 609

RESULT 14

US-09-949-016-122372/c

Sequence 122372, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 122372

LENGTH: 601

TYPE: DNA

ORGANISM: Human

US-09-949-016-122372

Query Match 3.4%; Score 79; DB 4; Length 601;

Best Local Similarity 89.5%; Pred. No. 4.7e-12;

Mismatches 85; Conservative 0; Indels 10; Gaps 0;

QY 425 AGAACCCAGAAATTCGAAGATGAAGCAAAAGAAACAATTCAGTCCCTTCTCCATGAAT 484

Db 268 AGAACCCAGAAATTCGAAGATGAAGCAAAAGAAACAATTCAGTCCCTTCTCCATGAAT 209

QY 485 GGTACACGACGAAACATGTTGAGGGGAGAAAGACTTGC 519

Db 208 GGTACACGACGAAACATGTTGAGGGGAGAAAGACTTGC 174

RESULT 15

US-08-232-463-14/c

Sequence 14, Application US/08232463

Patent No. 5670367

GENERAL INFORMATION:

APPLICANT: DORNER, F.

APPLICANT: SCHEIFLINGER, F.

APPLICANT: FALKNER, F. G.

TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner

STREET: 1800 Diagonal Road, Suite 500

CITY: Alexandria

STATE: VA

COUNTRY: USA

ZIP: 22313-0299

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/232,463

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/935,313

FILING DATE:

APPLICATION NUMBER: EP 91 114 300.6

FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: BENT, Stephen A.

REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 30472/114 IMMU

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)836-9300

TELEFAX: (703)683-4109

TELEX: 899149

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 7218 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

CLONE: pTZgpt-Fis

US-08-232-463-14

Query Match 3.3%; Score 77.2; DB 1; Length 7218;

Best Local Similarity 7.2%; Pred. No. 9.6e-11;

Mismatches 31; Conservative 238; Mismatches 161; Indels 0; Gaps 0;

QY 759 CATGATACAGATGATGAACATACCAAGTCTATCAGAGAAACAAGCATATATGAACCTCTA 818

Db 1490 CATCACTGTAATTACCTATCTATCAAGTAGTTAAAGAGATAGAGATTTGGTACRRR 1431

QY 819 GAAATGAAGGATAGAGAAATCACAGAAATCACTCCCTCCTGAGGATAATCTCTGTAGAA 878

Db 1430 RRR 1371

QY 879 GATTCAAGGTAATTGTAGAGAGATGAGCATTTTCTGTGGAGAACACAGCAAGTA 938



**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 14:35:53 ; Search time 9715 Seconds  
(without alignments)  
11591.346 Million cell updates/sec

Title: US-09-436-184-3  
Perfect score: 2324  
Sequence: 1 cggaccgtgcaatggccag.....cttgggaactctggagaga 2324

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466.

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hcg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2324	100.0	2324	6	AX146795 Sequence
2	2324	100.0	2324	9	S83325
3	2322	95.6	2452	6	CO871564
4	1766	76.0	2449	9	HSU03109
5	1164	50.1	1970	6	CO729666
6	895	38.5	2680	9	AF289489
7	684	29.4	2280	9	AF306765
8	640	27.5	2442	6	BD234762
9	640	27.5	2442	6	AR274038
10	640	27.5	2442	6	AR277619
11	640	27.5	2442	6	AR441744
12	640	27.5	2442	6	ARS32049
13	640	27.5	2442	6	AX369081
14	612	26.3	1080	9	BC015518
15	450	19.4	2761	9	BC025236
16	315	13.6	366	6	CO680735
17	297	12.8	297	6	CO731371
18	213	9.2	1608	9	BC066929
19	190	8.2	132868	2	AC021323 Homo sapi

20	190	8.2	158404	9	AC090094	AC090094 Homo sapi
21	172	7.4	855	9	AF224468	AF224468 Homo sapi
22	135	6.7	186676	9	AC067881	AC067881 Homo sapi
23	135	5.8	191	6	CO657781	CO657781 Sequence
24	130	5.6	126392	9	CR626894	CR626894 Human DNA
25	130	5.6	142265	9	AC091173	AC091173 Homo sapi
26	103	4.4	941	9	AF224469	AF224469 Homo sapi
27	103	4.4	1254	9	AF184241	AF184241 Homo sapi
28	62	2.7	2739	4	BOVASBYH	M31213 Bos taurus
29	60	2.6	60	6	CO541435	CO541435 Sequence
30	59	2.5	2208	6	AX771609	AX771609 Sequence
31	59	2.5	2208	6	HSATPBR	HSATPBR Human mRNA
32	56	2.4	177614	2	AC143562	AC143562 Macaca mu
33	53	2.3	1777	4	CFU38414	U38414 Canis famil
34	40	1.7	556	4	AF198966	AF198966 Oryctolag
35	35	1.5	4419	10	AF289486	AF289486 Mus muscu
36	35	1.5	6034	10	HO10058511	AF289215 Mus muscu
37	35	1.5	6629	10	AF289487	AF289487 Mus muscu
38	35	1.5	168569	10	AL671970	AL671970 Mouse DNA
39	32	1.4	524	10	AF289491	AF289491 Mus muscu
40	32	1.4	1763	10	BC061098	BC061098 Mus muscu
41	32	1.4	1951	10	HO10058502	AF289206 Mus muscu
42	32	1.4	2617	10	AF302653	AF302653 Mus muscu
43	32	1.4	2770	10	BC015281	BC015281 Mus muscu
44	32	1.4	2789	10	AF289488	AF289488 Mus muscu
45	32	1.4	198991	10	AL7713548	AL7713548 Mouse DNA
46	31	1.3	566	10	AF289492	AF289492 Mus muscu
47	31	1.3	919	10	AF221854	AF221854 Mus muscu
48	31	1.3	988	10	AF223414	AF223414 Mus muscu
49	31	1.3	1013	10	AF223413	AF223413 Mus muscu
50	31	1.3	1733	10	AF289490	AF289490 Mus muscu
51	31	1.3	2432	10	AF302655	AF302655 Mus muscu
52	31	1.3	2637	10	AF302654	AF302654 Mus muscu
53	30	1.3	158289	2	AC143465	AC143465 Macaca mu
54	29	1.2	234548	2	AC095635	AC095635 Rattus no
55	29	1.2	275547	2	AC128056	AC128056 Rattus no
56	27	1.2	27	6	CO871665	CO871665 Sequence
57	27	1.2	2013	10	HO10058505	AF289209 Mus muscu
58	27	1.2	252366	2	AC096178	AC096178 Rattus no
59	26	1.1	6629	10	HO10058507	AF289211 Mus muscu
60	26	1.1	12095	10	HO10058510	AF289214 Mus muscu
61	24	1.0	1519	10	AF016503	AF016503 Rattus no
62	24	1.0	1680	10	MMU19607	U19607 Mus musculu
63	24	1.0	2204	10	MMU16145	U16145 Mus musculu
64	24	1.0	2221	10	MUSRLK	L35268 Mus cookii
65	24	1.0	2342	10	MUSPTKRL18	D43963 Mus musculu
66	24	1.0	110000	2	AC013623_2	Continuation (3 of
67	24	1.0	110000	2	AC013623_3	Continuation (4 of
68	24	1.0	110000	8	CR382132_00	CR382132 Yarrowia
69	24	1.0	150010	5	BX663616	BX663616 Zebrafish
70	24	1.0	192578	5	AL732455	AL732455 Zebrafish
71	24	1.0	222930	2	AC099698	AC099698 Mus muscu
72	24	1.0	224556	10	AC102441	AC102441 Mus muscu
73	24	1.0	289405	2	AC090126	AC090126 Mus muscu
74	23	1.0	23	6	CO871636	CO871636 Sequence
75	23	1.0	595	11	G94809	G94809 S209P6230RC
76	23	1.0	2483	10	HO10058501	AF289205 Mus muscu
77	23	1.0	3146	10	AF026465	AF026465 Mus muscu
78	23	1.0	3198	10	BC053057	BC053057 Mus muscu
79	23	1.0	57919	2	AC017856	AC017856 Drosophila
80	23	1.0	115135	10	AC112161	AC112161 Mus muscu
81	23	1.0	120440	5	EX000467	EX000467 Zebrafish
82	23	1.0	163122	2	AC130052	AC130052 Rattus no
83	23	1.0	192136	10	AC133508	AC133508 Mus muscu
84	23	1.0	207761	3	AC093546	AC093546 Drosophila
85	23	1.0	210493	2	AC146886	AC146886 Callithiri
86	23	1.0	218311	2	CR354540	CR354540 Dario rer
87	23	1.0	220755	10	AC110235	AC110235 Mus muscu
88	23	1.0	228016	2	AC146674	AC146674 Callithiri
89	23	1.0	239875	2	AC126827	AC126827 Rattus no
90	23	1.0	272016	3	AE003539	AE003539 Drosophila
91	23	1.0	273077	2	AC094884	AC094884 Rattus no
92	22	0.9	855	6	CO608337	CO608337 Sequence



Db	1561	CCGAGATCTGGCACTGATGATGGAGATTTATTTCCACTTGGGGATGCCATGCAGA	1620
Qy	1621	GGTTGGGAAACAAGAGGCATATAAGTGTATGATGCTTGGGCAACAAGAGGACACTTTG	1680
Db	1621	GGTTGGGAAACAAGAGGCATATAAGTGTATGATGCTTGGGCAACAAGAGGACACTTTG	1680
Qy	1681	CATCTGTCTGGCAACGCTCACTCTCAATGTGAATGGAAGTGAAGACACACGCTTGGTGA	1740
Db	1681	CATCTGTCTGGCAACGCTCACTCTCAATGTGAATGGAAGTGAAGACACACGCTTGGTGA	1740
Qy	1741	CCCCAAAAGAAACGGGCTACACAGATGATGAATGCTTTTGAAGAAGAACTGGAATTA	1800
Db	1741	CCCCAAAAGAAACGGGCTACACAGATGATGAATGCTTTTGAAGAAGAACTGGAATTA	1800
Qy	1801	TCCGAGATGAAGGCTTCGAGTATGATGAAGCAAGGCTCTCTTCTGCTTGAAGATG	1860
Db	1801	TCCGAGATGAAGGCTTCGAGTATGATGAAGCAAGGCTCTCTTCTGCTTGAAGATG	1860
Qy	1861	AAACCTGAGGGAAGAAAGGGACTGAGCCAGCTTACGCTGTGGCAGCAGGAAGAGAA	1920
Db	1861	AAACCTGAGGGAAGAAAGGGACTGAGCCAGCTTACGCTGTGGCAGCAGGAAGAGAA	1920
Qy	1921	ATGAAATGCTCTCAAGGAGCTCTTAAACCTGTACTTACTAGAAAAAGTTCCCGAGA	1980
Db	1921	ATGAAATGCTCTCAAGGAGCTCTTAAACCTGTACTTACTAGAAAAAGTTCCCGAGA	1980
Qy	1981	CAACAGATGCAGAGGAGCAGATCAATATTCATCATGCAACCCGGGACTCAGCTGT	2040
Db	1981	CAACAGATGCAGAGGAGCAGATCAATATTCATCATGCAACCCGGGACTCAGCTGT	2040
Qy	2041	GGCGCACACAGGGCCCAAACTCAGGCTCCGAATGCACCTGGGCTTGGTATGCCA	2100
Db	2041	GGCGCACACAGGGCCCAAACTCAGGCTCCGAATGCACCTGGGCTTGGTATGCCA	2100
Qy	2101	AGGAAGCTCAAGATTCGATGTGCAACGAGACAGGACTTGGGAGGAGGCAAGGTGC	2160
Db	2101	AGGAAGCTCAAGATTCGATGTGCAACGAGACAGGACTTGGGAGGAGGCAAGGTGC	2160
Qy	2161	TCATCTTTGATGACTCCTTTGAGCAGGAGTATGGCAGGATGCTCATCTTCCGGCTGA	2220
Db	2161	TCATCTTTGATGACTCCTTTGAGCAGGAGTATGGCAGGATGCTCATCTTCCGGCTGA	2220
Qy	2221	TATTCATCGTGATGTGGCACTCGGAACCTGACACACAGAGAGCAGCCTTCCAG	2280
Db	2221	TATTCATCGTGATGTGGCACTCGGAACCTGACACACAGAGAGCAGCCTTCCAG	2280
Qy	2281	CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA	2324
Db	2281	CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA	2324
RESULT 2			
S83325			
LOCUS			
DEFINITION			
apartyl (asparaginyl) beta-hydroxylase [human, hepatoblastoma cell			
line Hep02, mRNA Partial, 2324 nt].			
S83325			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1 (bases 1 to 2324)			
Lavaresiere, L., Jia, S., Nishiyama, M., de la Monte, S., Stern, A.M.,			
Wands, J.R. and Friedman, P.A.			
Overexpression of human apartyl (asparaginyl) beta-hydroxylase in			
hepatocellular carcinoma and cholangiocarcinoma			
J. Clin. Invest. 98 (6), 1313-1323 (1996)			
98420598			
8823296			
PUBMED			
REMARK			
GenBank staff at the National Library of Medicine created this			
entry [NCBI gibseq 179962] from the original journal article.			

```
QY 541 GAGAACCAACAAGAGGATGATGAGTTCTTATGCGGCTGATGATGATGATGATG 600
DB 541 GAGAACCAACAAGAGGATGATGAGTTCTTATGCGGCTGATGATGATGATGATG 600
QY 601 AGACCTTGGAACTGGAAGTATCTCATGAAGAAACCGAGCATAGTATACCAGTGGAAAGAGA 660
DB 601 AGACCTTGGAACTGGAAGTATCTCATGAAGAAACCGAGCATAGTATACCAGTGGAAAGAGA 660
QY 661 CAGTTTCACAAGACTGTAATCAGGATATGGAAGAGATGATGTCGAGCAGGAAATCCAG 720
DB 661 CAGTTTCACAAGACTGTAATCAGGATATGGAAGAGATGATGTCGAGCAGGAAATCCAG 720
QY 721 ATTCCAGTGAACAGTATGAGATGAAAGATTTGCACCATGATACAGATGATGTAACAT 780
DB 721 ATTCCAGTGAACAGTATGAGATGAAAGATTTGCACCATGATACAGATGATGTAACAT 780
QY 781 ACCAAGTCTATGAGGAACACAGCATATATGAACCTCTAGAAATGAAAGGATGAAATCA 840
DB 781 ACCAAGTCTATGAGGAACACAGCATATATGAACCTCTAGAAATGAAAGGATGAAATCA 840
QY 841 CAGAAGTAACTGCTCCCTGAGGATATCTCTGTAGAAATTCACAGTAAATTTGTAAG 900
DB 841 CAGAAGTAACTGCTCCCTGAGGATATCTCTGTAGAAATTCACAGTAAATTTGTAAG 900
QY 901 AAGTAAGCATTTTCTCTGTGGAAGACAGCAGGAAGTACCAACAGAAACAAATAGAAAAA 960
DB 901 AAGTAAGCATTTTCTCTGTGGAAGACAGCAGGAAGTACCAACAGAAACAAATAGAAAAA 960
QY 961 CAGATGATCCAGAACCAAAAGTAAAGAAAGAAAGAGCCCTAACTTTTAAATTAAT 1020
DB 961 CAGATGATCCAGAACCAAAAGTAAAGAAAGAAAGAGCCCTAACTTTTAAATTAAT 1020
QY 1021 TTGATAAGCATTTTAAAGCTGAATCTGAGAAAGAAAGTAAAGAAAGAAAGGAAAGAA 1080
DB 1021 TTGATAAGCATTTTAAAGCTGAATCTGAGAAAGAAAGTAAAGAAAGAAAGGAAAGAA 1080
QY 1081 TTGAGGAAGCAGTGAATGATTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140
DB 1081 TTGAGGAAGCAGTGAATGATTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140
QY 1141 CAAGATATGGGAAGCGGAGTGTAGGATGATTTGGCTGAGAAGGAGAGTAAATGAGG 1200
DB 1141 CAAGATATGGGAAGCGGAGTGTAGGATGATTTGGCTGAGAAGGAGAGTAAATGAGG 1200
QY 1201 TGCTAGTGGAGCCATCGAGCCTACCAAGAGTGGCCAGCTACCTCATGTCCCTGCAG 1260
DB 1201 TGCTAGTGGAGCCATCGAGCCTACCAAGAGTGGCCAGCTACCTCATGTCCCTGCAG 1260
QY 1261 ACCTGCTGAAGCTGAGTTTGAAGCGCTGCTCAGACAGGCAACAAATTTCTAGTCAATGA 1320
DB 1261 ACCTGCTGAAGCTGAGTTTGAAGCGCTGCTCAGACAGGCAACAAATTTCTAGTCAATGA 1320
QY 1321 GAGGTTCCCTGCTTACCTTGGAGATGATTTCAACTATTTCCCAATGATCTCCCTTAA 1380
DB 1321 GAGGTTCCCTGCTTACCTTGGAGATGATTTCAACTATTTCCCAATGATCTCCCTTAA 1380
QY 1381 AAAATGACCTTGGCTGGGATACCTCTGATAGGAGATTAATGACAAATGCAAGAAAGTTT 1440
DB 1381 AAAATGACCTTGGCTGGGATACCTCTGATAGGAGATTAATGACAAATGCAAGAAAGTTT 1440
QY 1441 ATGAAGAGTGTGAGTGAACCTTAATGATGGCTTTGCTAAAGTCCATTTAGCTTCA 1500
DB 1441 ATGAAGAGTGTGAGTGAACCTTAATGATGGCTTTGCTAAAGTCCATTTAGCTTCA 1500
QY 1501 TCCCTGAAGGCACAGAACAAATGCTGAGAGATCCATATTTAAAGGAAGAAATAGAT 1560
DB 1501 TCCCTGAAGGCACAGAACAAATGCTGAGAGATCCATATTTAAAGGAAGAAATAGAT 1560
QY 1561 CCGGAGATCTCGGCACTGATGATGGGAGATTTTATTTCCACTGGGGGATGCCATGAGA 1620
DB 1561 CCGGAGATCTCGGCACTGATGATGGGAGATTTTATTTCCACTGGGGGATGCCATGAGA 1620
QY 1621 GGGTTGGGAACAAGAGGCATATTAAGTGGTATGAGCTTGGGCACAGAGAGGACATTTG 1680
```

```
DB 1621 GGGTTGGGAACAAGAGGCATATTAAGTGGTATGAGCTTGGGCACAGAGAGACACTTTG 1680
QY 1681 CATCTGTCTGGCAACGCTCACTCTACAAATGTGAATGAGTGAAGACACAGCCTTGGTGA 1740
DB 1681 CATCTGTCTGGCAACGCTCACTCTACAAATGTGAATGAGTGAAGACACAGCCTTGGTGA 1740
QY 1741 CCCCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGCTTTTAGAAAAGAACTGGAAGTTAA 1800
DB 1741 CCCCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGCTTTTAGAAAAGAACTGGAAGTTAA 1800
QY 1801 TCCGAGATGAAGGCTTGGCAGTATGATGATTAAGCCAAAGGCTCTTCTGCTGCTGAGGATG 1860
DB 1801 TCCGAGATGAAGGCTTGGCAGTATGATGATTAAGCCAAAGGCTCTTCTGCTGCTGAGGATG 1860
QY 1861 AAAAAGCTGAGGAAAGAAAGGGGACTGAGCCAGTCTCAAGCTGTGGCAGCAAGGAAGAA 1920
DB 1861 AAAAAGCTGAGGAAAGAAAGGGGACTGAGCCAGTCTCAAGCTGTGGCAGCAAGGAAGAA 1920
QY 1921 ATGAAAATGCTGCAAGAGGAGCTCTTAAACCTGTACTTACTAGAAAAGTTCCCGGAGA 1980
DB 1921 ATGAAAATGCTGCAAGAGGAGCTCTTAAACCTGTACTTACTAGAAAAGTTCCCGGAGA 1980
QY 1981 CAACAGATGCAAGAGGAGGACAGATCAATATTTCCATCATGCAACCCCGGACTCACGTGT 2040
DB 1981 CAACAGATGCAAGAGGAGGACAGATCAATATTTCCATCATGCAACCCCGGACTCACGTGT 2040
QY 2041 GGCGGACACACAGGCGCCCAACAACTGCAGGCTCCGAATGCACCTGGGCTTGGTGAATCCCA 2100
DB 2041 GGCGGACACACAGGCGCCCAACAACTGCAGGCTCCGAATGCACCTGGGCTTGGTGAATCCCA 2100
QY 2101 AGAAGGCTGCAAGATTCGATGTGCAACAGACAGGACCTGGGAGGAAGCAAGGTGC 2160
DB 2101 AGAAGGCTGCAAGATTCGATGTGCAACAGGACCTGGGAGGAAGCAAGGTGC 2160
QY 2161 TCATCTTGTATGATCTCTTTTGACAGAGTATGCGAGATGCGCTCATCTTTCCGCTGA 2220
DB 2161 TCATCTTGTATGATCTCTTTTGACAGAGTATGCGAGATGCGCTCATCTTTCCGCTGA 2220
QY 2221 TATTCATCTGTGATGTGTGCAATCCGAACTGACACACAGCAGACCGAGCCTTCCAG 2280
DB 2221 TATTCATCTGTGATGTGTGCAATCCGAACTGACACACAGCAGACCGAGCCTTCCAG 2280
QY 2281 CAATTTAGCATGAATTCATGCAAGCTTGGAAACTCTGGAGAGA 2324
DB 2281 CAATTTAGCATGAATTCATGCAAGCTTGGAAACTCTGGAGAGA 2324
```

## RESULT 3

LOCUS CQ871564 2452 bp DNA linear PAT 27-SEP-2004  
DEFINITION Sequence 55 from Patent WO2004078783.

ACCESSION CQ871564

VERSION CQ871564.1 GI:52745624

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 Murphy, F.J., Sheehan, D.E., Keating, K.E., Hayes, I.E. and Seera, L.E.

Enzymes involved in apoptosis

Patent: WO 2004078783-A 55 16-SEP-2004;

Elrx Therapeutics Ltd (IE)

Location/Qualifiers

1..2452

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

## FEATURES

source

ORIGIN

Query Match 95.6%; Score 2222; DB 6; Length 2452;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2322; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	CGGACCGTGCATATGCGCCAGCGTAAAGATGCCAAGAGCAGCGCAACAGCAGCAGCGG	60
Db	67	CGGACCGTGCATATGCGCCAGCGTAAAGATGCCAAGAGCAGCGCAACAGCAGCAGCGG	126
Qy	61	GCTCCGCGCAGCGTACGACGAGTGGGGCAGCAGCAGCGCCCGGGCCCGGAGAGACAA	120
Db	127	GCTCCGCGCAGCGTACGACGAGTGGGGCAGCAGCAGCGCCCGGGCCCGGAGAGACAA	186
Qy	121	AGCATGGAGGACACAAAGATGGGAGGAAAGCGCGACTCTCGGGAACCTTCACTTTCACGT	180
Db	187	AGCATGGAGGACACAAAGATGGGAGGAAAGCGCGACTCTCGGGAACCTTCACTTTCACGT	246
Qy	181	GGTTTATGGTGATTCGATTGCTGGGCGCTGGGACATCTGTAGCTGTCTGTTGGTTTGATC	240
Db	247	GGTTTATGGTGATTCGATTGCTGGGCGCTGGGACATCTGTAGCTGTCTGTTGGTTTGATC	306
Qy	241	TTGTTGACTATGAGGAAGTCTAGGAAACTAGGAATCTATGATGCTGATGCTGATGGAG	300
Db	307	TTGTTGACTATGAGGAAGTCTAGGAAACTAGGAATCTATGATGCTGATGCTGATGGAG	366
Qy	301	ATTTTGATGTGATGATGCCAAAGTTTATAGGACTTAAAGAGAGATCTACTTTCAGAGC	360
Db	367	ATTTTGATGTGATGATGCCAAAGTTTATAGGACTTAAAGAGAGATCTACTTTCAGAGC	426
Qy	361	CAGCAGTCCCGCAGAGAGCGCTGAGCCACACACTGAGCCCGGAGCAGGTTCCGTGTG	420
Db	427	CAGCAGTCCCGCAGAGAGCGCTGAGCCACACACTGAGCCCGGAGCAGGTTCCGTGTG	486
Qy	421	AGCGAGAACCCAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCATG	480
Db	487	AGCGAGAACCCAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCATG	546
Qy	481	AAATGGTACAGCAGACACATGTTGAGGGAGAGACTTGGCAACAGAGAGTGGACCCACAG	540
Db	547	AAATGGTACAGCAGACACATGTTGAGGGAGAGACTTGGCAACAGAGAGTGGACCCACAG	606
Qy	541	GAGAACCAACACAGAGGATGATGATTTCTTATGGCGACTGATGTAGATGATAGATTG	600
Db	607	GAGAACCAACACAGAGGATGATGATTTCTTATGGCGACTGATGTAGATGATAGATTG	666
Qy	601	AGACCTGGAACTCTGAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGAGAGA	660
Db	667	AGACCTGGAACTCTGAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGAGAGA	726
Qy	661	CAGTTTCAACAAGCTGTAATCAGGATATGGAAGATGATGTTGAGCAGGAAATCCAG	720
Db	727	CAGTTTCAACAAGCTGTAATCAGGATATGGAAGATGATGTTGAGCAGGAAATCCAG	786
Qy	721	ATTCAGTGAACAGTAGTAGAAGATGGAAGATTCGACCATGATACAGATGATGTAACAT	780
Db	787	ATTCAGTGAACAGTAGTAGAAGATGGAAGATTCGACCATGATACAGATGATGTAACAT	846
Qy	781	ACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA	840
Db	847	ACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA	906
Qy	841	CAGAAGTAACTGCTCCCGCTGAGGATAATCTGTAGAGATTCACAGGTAATGTAGAG	900
Db	907	CAGAAGTAACTGCTCCCGCTGAGGATAATCTGTAGAGATTCACAGGTAATGTAGAG	966
Qy	901	AGTAAGCATTTTCTGTTGGAGACAGCAGGAGTACCCACGAAACAAATAGAAAAA	960
Db	967	AGTAAGCATTTTCTGTTGGAGACAGCAGGAGTACCCACGAAACAAATAGAAAAA	1026
Qy	961	CAGATGATCCAGAACAAAGCAAAAGTTTAAAGAAAGAGCCCTTAAATAAAT	1020
Db	1027	CAGATGATCCAGAACAAAGCAAAAGTTTAAAGAAAGAGCCCTTAAATAAAT	1086
Qy	1021	TTGATAAGACTATTAAAGCTGAACCTTGATGCTGCAGAAAAAATCCCGTAAAGGGGAAAA	1080
Db	1087	TTGATAAGACTATTAAAGCTGAACCTTGATGCTGCAGAAAAAATCCCGTAAAGGGGAAAA	1146

Qy	1081	TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGCGAAATACCTCAGAGTCCACGAG	1140
Db	1147	TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGCGAAATACCTCAGAGTCCACGAG	1206
Qy	1141	CAAGATATGGGAAGCGCAGTGTAGGATGATTTGGCTGAGAAGAGAGGAAGTAATGAGG	1200
Db	1207	CAAGATATGGGAAGCGCAGTGTAGGATGATTTGGCTGAGAAGAGAGGAAGTAATGAGG	1266
Qy	1201	TGCTACGTGGAGCATCGAGACCTCAAGAGGTGGCCAGCTACCTGATGTCCCTGCGAG	1260
Db	1267	TGCTACGTGGAGCATCGAGACCTCAAGAGGTGGCCAGCTACCTGATGTCCCTGCGAG	1326
Qy	1261	ACCTGCTGAAGCTGAGTTTGAAGCGTGGCTCAGACAGGCAACAATTTCTAGGTCATATGA	1320
Db	1327	ACCTGCTGAAGCTGAGTTTGAAGCGTGGCTCAGACAGGCAACAATTTCTAGGTCATATGA	1386
Qy	1321	GAGGTTCCCTGCTTACCTGCGAGATAGTTCAAATATTTCCCAATGATACCTTCTTAA	1380
Db	1387	GAGGTTCCCTGCTTACCTGCGAGATAGTTCAAATATTTCCCAATGATACCTTCTTAA	1446
Qy	1381	AAAAATGACCTTGGCGTGGATACCTCTGATAGGAGATAATGACAAATGCAAGAAAGTTT	1440
Db	1447	AAAAATGACCTTGGCGTGGATACCTCTGATAGGAGATAATGACAAATGCAAGAAAGTTT	1506
Qy	1441	ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGGCTTGTAAAGTCCATTATGGCTTCA	1500
Db	1507	ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGGCTTGTAAAGTCCATTATGGCTTCA	1566
Qy	1501	TCCTGAAGGACAGAAACAAATTTCTGAGAGCATCCCATATTTTAAAGGAAGGAATAGAA	1560
Db	1567	TCCTGAAGGACAGAAACAAATTTCTGAGAGCATCCCATATTTTAAAGGAAGGAATAGAA	1626
Qy	1561	CCGAGATCTCTGGCCTGATGATGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGA	1620
Db	1627	CCGAGATCTCTGGCCTGATGATGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGA	1686
Qy	1621	GGTGTGGGAAACAAAGAGGACATATAAGTGTATGAGCTTGGGCACAGAGAGGACATTG	1680
Db	1687	GGTGTGGGAAACAAAGAGGACATATAAGTGTATGAGCTTGGGCACAGAGAGGACATTG	1746
Qy	1681	CATCTGTCTGCAACCTCCTACATGATGATGATGATGATGATGATGATGATGATGATG	1740
Db	1747	CATCTGTCTGCAACCTCCTACATGATGATGATGATGATGATGATGATGATGATGATG	1806
Qy	1741	CCCCAAAGAAACCGGCTACACAGAGTTAGTAAAGTCTTTTAAAGAAACCTGGAAGTTAA	1800
Db	1807	CCCCAAAGAAACCGGCTACACAGAGTTAGTAAAGTCTTTTAAAGAAACCTGGAAGTTAA	1866
Qy	1801	TCCGAGATGAAGGCTTTCAGTGTGATGATGATGATGATGATGATGATGATGATGATG	1860
Db	1867	TCCGAGATGAAGGCTTTCAGTGTGATGATGATGATGATGATGATGATGATGATGATG	1926
Qy	1861	AAAACTGTGGGAAACCGGCTGAGCAGTTTCAAGTGTGGGAGGAGGAGGAGGAGGAG	1920
Db	1927	AAAACTGTGGGAAACCGGCTGAGCAGTTTCAAGTGTGGGAGGAGGAGGAGGAGGAG	1986
Qy	1921	ATGAAATGCTGCAAGAGGCTTCTAAACCTGTATCTTTAAGAAAAGTTTCCCGGAGA	1980
Db	1987	ATGAAATGCTGCAAGAGGCTTCTAAACCTGTATCTTTAAGAAAAGTTTCCCGGAGA	2046
Qy	1981	CAACAGGATGCAAGAGGAGCAGATCAATATTTCCATCATGCAACCCCGGAGCTCACGTGT	2040
Db	2047	CAACAGGATGCAAGAGGAGCAGATCAATATTTCCATCATGCAACCCCGGAGCTCACGTGT	2106
Qy	2041	GGCGGACACAGCGGCCCAAACTGCGAGGTCCGAATGCACTGGCTTGGTGTATCCCA	2100
Db	2107	GGCGGACACAGCGGCCCAAACTGCGAGGTCCGAATGCACTGGCTTGGTGTATCCCA	2166
Qy	2101	AGGAAGGCTGCAAGATTCGATGTGCCAACGAGACCGAGGACCTGGGAGGAGGAGGTGC	2160
Db	2167	AGGAAGGCTGCAAGATTCGATGTGCCAACGAGACCGAGGACCTGGGAGGAGGAGGTGC	2226

QY 2161 TCATCTTTGATGACTCCTTTGAGCAGGAGGTATGGCAGGATGCTCATCTTCCGGCTGA 2220  
DB 2227 TCATCTTTGATGACTCCTTTGAGCAGGAGGTATGGCAGGATGCTCATCTTCCGGCTGA 2286  
QY 2221 TATTCATCGTGGATGTGTGGCATCCGGAATCGACACACACAGCAGCAGCAGCCTTCCAG 2280  
DB 2287 TATTCATCGTGGATGTGTGGCATCCGGAATCGACACACACAGCAGCAGCAGCCTTCCAG 2346  
QY 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2324  
DB 2347 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2390

RESULT 4  
HSU03109 2449 bp mRNA linear PRI 30-NOV-1995  
LOCUS Human aspartyl beta-hydroxylase mRNA, complete cds.

DEFINITION  
ACCESSION U03109  
VERSION U03109.1 GI:458031  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS Koriath, F., Gieffers, C. and Frey, J.  
TITLE Cloning and characterization of the human gene encoding aspartyl  
beta-hydroxylase

JOURNAL  
MEDLINE Gene 150 (2), 395-399 (1994)  
PUBMED 7821814

REFERENCE  
AUTHORS Koriath, F.  
TITLE Direct Submission  
SUBMITTED (03-NOV-1993) Koriath F., Fakultät fuer Chemie-Biochemie  
II, Universitaet Bielefeld, Universitaetsstrasse 25, Bielefeld,  
33615, Germany

FEATURES  
source  
1..2449  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="As-5"  
/cell\_line="mg63"  
/cell\_type="osteosarcoma"  
/clone\_lib="MG63-ZAP"  
1..77  
78..2351  
/function="hydroxylation of aspartyl and asparaginyl  
residues"  
/codon\_start=1  
/product="aspartyl beta-hydroxylase"  
/protein\_id="AAA82108.1"  
/db\_xref="GI:458032"  
/translation="MAQRNKAQSSGSSGSGSTAGSSSPQARRETHKGGHNG  
RKGLSGTSFTFWFVWVIALGLVSVVWVDFVLYEVLGKLIYDADGDFDVID  
AKVLLGLKERSSEPAVPEEAEPTPEEQVPEAEQNIIDEAEQIQLLHBMVH  
AEHVREDLQEDQDPTGPEQEDDEFMATDVRFTLEPEVSHETESHVHEETI  
SQCNODMEMMSSEOPDSSSPVVEERLRLHDDVTYVVEQVAVPLENEGIEI  
TEVAPENPDVDSQVIVVEVSIIPVEEQVPEPTNPKTDTPQAKVKKPKLL  
NKFDKTIKALDAAEKRRKRIEAVNAFKELVKYPQSPRARGKACQEDDLAEKR  
RSNEVLGAIETYQEVASLPDPADLLLSLKRSDRQOFLGHRGSLTLQLVOLF  
PNDTSILKNDLGIYDNDNAKVEEVLSTFNDGPAKVHYGFIKAQNKIAESI  
PYLKEGLSDQDPTGDRFVFLHGDAMQVGNKEAYKVELGKRGHGFASVWQRLIN  
VNLKAQCGPKSTGTOLVKSLEARNKLI RDEGLAVMDKAKGLFLPDENLRKGDW  
SQFTLQQQRNENACKGPKCTLLEKEPETTGCRGOIKYSIMHPGTHVWPHGTPT  
NCLRHLGLVLPKRGCKINCANETKEEGKVLIFDDSFHEVWQDASSPRLIFIVD  
VWHPELTPQOORSLPAI"  
2352..2449

5'UTR  
CDS

3'UTR  
ORIGIN

Query Match 76.08; Score 1766; DB 9; Length 2449;  
Best Local Similarity 99.74; Pred. No. 0;

Matches 2316; Conservative 0; Mismatches 5; Indels 3; Gaps 2;  
QY 1 CGGACCGTGCATGGCCAGCGCTAGAAATGCCAAGCAGCAGCGGCAACAGCAGCAGCAGCG 60  
DB 67 CGGACCGTGCATGGCCAGCGCTAGAAATGCCAAGCAGCAGCGGCAACAGCAGCAGCAGCG 126  
QY 61 GCTCCGGCAGCGGTAGCAGCAGTGGGGCAGCAGCAGCGCCCGGGCCCGGAGAGAGACAA 120  
DB 127 GCTCCGGCAGCGGTAGCAGCAGTGGGGCAGCAGCAGCGCCCGGGCCCGGAGAGAGACAA 186  
QY 121 AGCATGGAGCACAGAAATGGGAAAGCGGACTCTCGGGAACCTTCATCTTTCACGT 180  
DB 187 AGCATGGAGCACAGAAATGGGAAAGCGGACTCTCAGAACTTCATCTTTCACGT 246  
QY 181 GGTTCATGTGATTCATTCCTGGGCTCTGGACATCTCTAGCTGTCTGTTGGTTTGTATC 240  
DB 247 GGTTCATGTGATTCATTCCTGGGCTCTGGACATCTCTAGCTGTCTGTTGGTTTGTATC 306  
QY 241 TTCTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGTTGATGGAG 300  
DB 307 TTCTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGTTGATGGAG 366  
QY 301 ATTTTGATGTGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTTACTTTCAGAGC 360  
DB 367 ATTTTGATGTGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTTACTTTCAGAGC 426  
QY 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGGAGTTCCTGTGG 420  
DB 427 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGGAGTTCCTGTGG 486  
QY 421 AGCAGAACCCCAAGATATCGAAGATGAAAGAAAGAACTTCTAGTCCCTTCTCCATG 480  
DB 487 AGCAGAACCCCAAGATATCGAAGATGAAAGAAAGAACTTCTAGTCCCTTCTCCATG 546  
QY 481 AATGGTGTACGCGCAGAACATGTTGAGGAGAGAGACTTGCACAAAGAGATGAGCCACAG 540  
DB 547 AATGGTGTACGCGCAGAACATGTTGAGGAGAGAGACTTGCACAAAGAGATGAGCCACAG 606  
QY 541 GAGAACCAACAGAGAGATGATGATTTCTTATGGGACTGATGATGATGATGATGATGATG 600  
DB 607 GAGAACCAACAGAGAGATGATGATTTCTTATGGGACTGATGATGATGATGATGATGATG 666  
QY 601 AGACCTTGGAACCTGGAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGAAGAGA 660  
DB 667 AGACCTTGGAACCTGGAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGAAGAGA 726  
QY 661 CAGTTTCAACAAGCTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAGAAATCCAG 720  
DB 727 CAGTTTCAACAAGCTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAGAAATCCAG 786  
QY 721 ATTCCAGTGAACAGTATGAGAGATGAAAGATGCAACCATGATGATGATGATGATGATGATG 780  
DB 787 ATTCCAGTGAACAGTATGAGAGATGAAAGATGCAACCATGATGATGATGATGATGATGATG 846  
QY 781 ACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAAAATGAAGGATAGAAATCA 840  
DB 847 ACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAAAATGAAGGATAGAAATCA 906  
QY 841 CAGAAGTAATCTGCTCCCTGAGGATTAATCTCTGAGGATTAATCTCTGAGGATTAATCTGAGG 900  
DB 907 CAGAAGTAATCTGCTCCCTGAGGATTAATCTCTGAGGATTAATCTCTGAGGATTAATCTGAGG 966  
QY 901 AAGTAAGCATTTTCTCTGAGGAAACAGCAGGAGGATGATGATGATGATGATGATGATGATG 960  
DB 967 AAGTAAGCATTTTCTCTGAGGAAACAGCAGGAGGATGATGATGATGATGATGATGATGATG 1026  
QY 961 CAGATGATCCAGAAACAAAGCTTAAAGAAAAAGAGCCCTAAACTTTTAAATAAAT 1020  
DB 1027 CAGATGATCCAGAAACAAAGCTTAAAGAAAAAGAGCCCTAAACTTTTAAATAAAT 1086  
QY 1021 TTGATAAGCATTTTAAAGCTGAACTTGTGATGCTGAGAAAAAACTCCGTAAGAGGGGAAAAA 1080  
DB 1087 TTGATAAGCATTTTAAAGCTGAACTTGTGATGCTGAGAAAAAACTCCGTAAGAGGGGAAAAA 1146













## RESULT 10

AR277619  
LOCUS  
DEFINITION  
Accession  
AR277619.1 GI:29711268  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Unknown.  
Unclassified.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source

Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S.,  
Carter, D., Retter, M.W., Mannion, J., Fan, L. and Wang, A.  
Compositions and methods for the therapy and diagnosis of lung  
cancer  
Patent: US 6509448-A 1791 21-JAN-2003;  
Location/Qualifiers  
1..2442  
/organism="unknown"  
/mol\_type="genomic DNA"

## ORIGIN

Query Match 27.5%; Score 640; DB 6; Length 2442;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 162 GGAATTCATCTTCACGTGGTTTATGGTGGTTCATGTCGCGGCTCGGACATCTGTA 221  
DB 46 GGAATTCATCTTCACGTGGTTTATGGTGGTTCATGTCGCGGCTCGGACATCTGTA 105  
QY 222 GCTGTCGTTTGGTTTGGTTCATCTGTCGACTATGAGGAAGTCTTAGGAAACTAGGAATCTAT 281  
DB 106 GCTGTCGTTTGGTTTGGTTCATCTGTCGACTATGAGGAAGTCTTAGGAAACTAGGAATCTAT 165  
QY 282 GATGTCGATGTCGATGAGATTTTGGTGGTATGTCGCAAGTCTTATAGCACTTAA 341  
DB 166 GATGTCGATGTCGATGAGATTTTGGTGGTATGTCGCAAGTCTTATAGCACTTAA 225  
QY 342 GAGAGTCTACTTCAGAGCCAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCC 401  
DB 226 GAGAGTCTACTTCAGAGCCAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCC 285  
QY 402 GAGGAGCAGGTTCTCTGAGGCGAGAACCCAGATATCGAAGATGAGCAAGAACAA 461  
DB 286 GAGGAGCAGGTTCTCTGAGGCGAGAACCCAGATATCGAAGATGAGCAAGAACAA 345  
QY 462 ATTCAAGTCCCTTCTCCATGAAATGGTACACGAGAACATGTTGAGGAGAGACTTGC 521  
DB 346 ATTCAAGTCCCTTCTCCATGAAATGGTACACGAGAACATGTTGAGGAGAGACTTGC 405  
QY 522 CAAGAGATGGAACCCACAGGAGAACCAACAAAGAGGATGATGATTTCTTATGCGGACT 581  
DB 406 CAAGAGATGGAACCCACAGGAGAACCAACAAAGAGGATGATGATTTCTTATGCGGACT 465  
QY 582 GATGTAGATGATGATTTGAGACCTTGGAACTTGAAGTATCTCATGAAGAAACCGAGCAT 641  
DB 466 GATGTAGATGATGATTTGAGACCTTGGAACTTGAAGTATCTCATGAAGAAACCGAGCAT 525  
QY 642 AGTTACCACTGPGGAAGAGACAGTTTCAAGAATCTGATATCAGGATATGGAAGAGATGATG 701  
DB 526 AGTTACCACTGPGGAAGAGACAGTTTCAAGAATCTGATATCAGGATATGGAAGAGATGATG 585  
QY 702 TCTGAGCAGGAGAAATCCAGATTCAGTGAAACAGTAGTAGGATGAAAGATGCAACCAT 761  
DB 586 TCTGAGCAGGAGAAATCCAGATTCAGTGAAACAGTAGTAGGATGAAAGATGCAACCAT 645  
QY 762 GATACAGATGATGTAACATACCAAGTCTTATGAGGAACCAAGCAGTATATGAACCTCTAGAA 821  
DB 646 GATACAGATGATGTAACATACCAAGTCTTATGAGGAACCAAGCAGTATATGAACCTCTAGAA 705  
QY 822 AATGAAGGGATAGAAATCACAAGGTAACCTG 852  
DB 706 AATGAAGGGATAGAAATCACAAGGTAACCTG 736

## RESULT 11

AR441744  
LOCUS  
DEFINITION  
Accession  
AR441744  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Unknown.  
Unclassified.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source

Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S.,  
Carter, D., Retter, M.W., Mannion, J., Fan, L.  
Compositions and methods for the therapy and diagnosis of lung  
cancer  
Patent: US 6667154-A 1791 23-DEC-2003;  
Location/Qualifiers  
1..2442  
/organism="unknown"  
/mol\_type="genomic DNA"

## ORIGIN

Query Match 27.5%; Score 640; DB 6; Length 2442;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 162 GGAATTCATCTTCACGTGGTTTATGGTGGTTCATGTCGCGGCTCGGACATCTGTA 221  
DB 46 GGAATTCATCTTCACGTGGTTTATGGTGGTTCATGTCGCGGCTCGGACATCTGTA 105  
QY 222 GCTGTCGTTTGGTTTGGTTCATCTGTCGACTATGAGGAAGTCTTAGGAAACTAGGAATCTAT 281  
DB 106 GCTGTCGTTTGGTTTGGTTCATCTGTCGACTATGAGGAAGTCTTAGGAAACTAGGAATCTAT 165  
QY 282 GATGTCGATGTCGATGAGATTTTGGTGGTATGTCGCAAGTCTTATAGCACTTAA 341  
DB 166 GATGTCGATGTCGATGAGATTTTGGTGGTATGTCGCAAGTCTTATAGCACTTAA 225  
QY 342 GAGAGTCTACTTCAGAGCCAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCC 401  
DB 226 GAGAGTCTACTTCAGAGCCAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCC 285  
QY 402 GAGGAGCAGGTTCTCTGAGGCGAGAACCCAGATATCGAAGATGAGCAAGAACAA 461  
DB 286 GAGGAGCAGGTTCTCTGAGGCGAGAACCCAGATATCGAAGATGAGCAAGAACAA 345  
QY 462 ATTCAAGTCCCTTCTCCATGAAATGGTACACGAGAACATGTTGAGGAGAGACTTGC 521  
DB 346 ATTCAAGTCCCTTCTCCATGAAATGGTACACGAGAACATGTTGAGGAGAGACTTGC 405  
QY 522 CAAGAGATGGAACCCACAGGAGAACCAACAAAGAGGATGATGATTTCTTATGCGGACT 581  
DB 406 CAAGAGATGGAACCCACAGGAGAACCAACAAAGAGGATGATGATTTCTTATGCGGACT 465  
QY 582 GATGTAGATGATGATTTGAGACCTTGGAACTTGAAGTATCTCATGAAGAAACCGAGCAT 641  
DB 466 GATGTAGATGATGATTTGAGACCTTGGAACTTGAAGTATCTCATGAAGAAACCGAGCAT 525  
QY 642 AGTTACCACTGPGGAAGAGACAGTTTCAAGAATCTGATATCAGGATATGGAAGAGATGATG 701  
DB 526 AGTTACCACTGPGGAAGAGACAGTTTCAAGAATCTGATATCAGGATATGGAAGAGATGATG 585  
QY 702 TCTGAGCAGGAGAAATCCAGATTCAGTGAAACAGTAGTAGGATGAAAGATGCAACCAT 761  
DB 586 TCTGAGCAGGAGAAATCCAGATTCAGTGAAACAGTAGTAGGATGAAAGATGCAACCAT 645  
QY 762 GATACAGATGATGTAACATACCAAGTCTTATGAGGAACCAAGCAGTATATGAACCTCTAGAA 821  
DB 646 GATACAGATGATGTAACATACCAAGTCTTATGAGGAACCAAGCAGTATATGAACCTCTAGAA 705  
QY 822 AATGAAGGGATAGAAATCACAAGGTAACCTG 852

[illegible]







```

Db      667 AAATGTTACACGCAAAATG 687
|||||
RESULT 16
LOCUS   CQ680735
DEFINITION Sequence 25661 from Patent WO02070737.
ACCESSION CQ680735
VERSION   CQ680735.1 GI:42196629
KEYWORDS Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS  Liaw, C.C., Marshall, W.E. and Zhang, H.
TITLE    Compositions and methods relating to osteoarthritis
JOURNAL  Patent: WO 02070737-A 25661 12-SEP-2002;
          ChondroGene Inc. (CA)
FEATURES
source   1..366
          /organism="Homo sapiens"
          /mol_type="unassigned DNA"
          /db_xref="taxon:9606"

ORIGIN
Query Match      13.6%; Score 315; DB 6; Length 366;
Best Local Similarity 99.7%; Pred. No. 3.2e-161; Indels 0; Gaps 0;
Matches 365; Conservative 0; Mismatches 1;

Qy 191 GATTCGATGCTGGGCGTCTGGACATCTGTAGCTGCTGGTTGGTTGATCTTGGTCACTA 250
Db 1 GATTCGATGCTGGGCGTCTGGACATCTGTAGCTGCTGGTTGGTTGATCTTGGTCACTA 60

Qy 251 TCAGGAAGTTCTAGGAAACTAGGAATCTATGATCTGATGCTGATGGAGATTTGATGT 310
Db 61 TCAGGAAGTTCTAGGAAACTAGGAATCTATGATCTGATGCTGATGGAGATTTGATGT 120

Qy 311 GGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGCAGCAGTCCC 370
Db 121 GGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGCAGCAGTCCC 180

Qy 371 GCCAAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTGTGGAGCAGAAC 430
Db 181 GCCAAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTGTGGAGGAGAAC 240

Qy 431 CCAGAAATATCGAAGATCAAGCAAAAGAAATTCAGTCCCTTCTCCATGAATGGTACA 490
Db 241 CCAGAAATATCGAAGATCAAGCAAAAGAAATTCAGTCCCTTCTCCATGAATGGTACA 300

Qy 491 CCAGAAATATGTTGAGGAGAGACTTGCACAAAGAGATGGACCCACAGAGAACCA 550
Db 301 CCAGAAATATGTTGAGGAGAGACTTGCACAAAGAGATGGACCCACAGAGAACCA 360

Qy 551 ACACGA 556
Db 361 ACACGA 366

RESULT 17
LOCUS   CQ731371
DEFINITION Sequence 17305 from Patent WO02068579.
ACCESSION CQ731371
VERSION   CQ731371.1 GI:42307776
KEYWORDS Homo sapiens (human)
SOURCE    Homo sapiens
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS  Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.

```

---

```

TITLE      Kites, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
JOURNAL    Patent: WO 02068579-A 17305 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES
source     1..297
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Query Match      12.8%; Score 297; DB 6; Length 297;
Best Local Similarity 100.0%; Pred. No. 2.7e-151; Indels 0; Gaps 0;
Matches 297; Conservative 0; Mismatches 0;

Qy 334 GACTTAAAGAGAGATCTACTTCAGACCCGAGAGTCCCGCAGAGAGGCTTGAGCCACACA 393
Db 1 GACTTAAAGAGAGATCTACTTCAGACCCGAGAGTCCCGCAGAGAGGCTTGAGCCACACA 60

Qy 394 CTGAGCCCGAGGAGCAGGTTCTCTGTGGAGCGAGAACCCCAAGATATCGAAGTGAAGCAA 453
Db 61 CTGAGCCCGAGGAGCAGGTTCTCTGTGGAGCGAGAACCCCAAGATATCGAAGTGAAGCAA 120

Qy 454 AAGAACAAATTCAGTCCCTTCTCCATGAATGGTACAGCAGAGACATGTTGAGGAGAG 513
Db 121 AAGAACAAATTCAGTCCCTTCTCCATGAATGGTACAGCAGAGACATGTTGAGGAGAG 180

Qy 514 ACTTGACACAGAGAGATGGACCCACAGGAGACACACACAGAGAGATGATGATGTTCTTA 573
Db 181 ACTTGACACAGAGAGATGGACCCACAGGAGACACACACAGAGAGATGATGATGTTCTTA 240

Qy 574 TGGCGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 630
Db 241 TGGCGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 297

RESULT 18
LOCUS   BC066929
DEFINITION Homo sapiens cDNA clone MGC:87227 IMAGE:5264813, complete cds.
ACCESSION BC066929
VERSION   BC066929.1 GI:44890460
KEYWORDS MGC.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS  Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
          Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
          Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
          Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
          Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
          Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
          Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
          Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
          Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
          McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
          Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
          Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
          Fahey, J., Helton, E., Kettman, M., Madan, A., Young, A.C., Shavchenko, Y.,
          Sanchez, A., Whiting, R.W., Madan, A., Young, A.C., Shavchenko, Y.,
          Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
          Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
          Butterfield, Y.S., Krzywinski, M.I., Skalka, U., Smaluk, D.E.,
          Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
          Generation and initial analysis of more than 15,000 full-length
          human and mouse cDNA sequences
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED    12477932
REFERENCE 2 (bases 1 to 1608)
AUTHORS  Strausberg, R.

```

**TITLE**  
JOURNAL

**REMARK**  
**COMMENT**

Direct Submission  
Submitted (01-MAR-2004) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki  
Toshiyuki and Piero Carninci (RIKEN)  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
Dickson, M., Schmutz, J., Grilwood, J., Rodriguez, A., and Myers,  
R. M.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAP Plate: 167 Row: k Column: 17  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 14589865.  
Location/Qualifiers  
1. 1608  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:87227 IMAGE:5264813"  
/tissue type="Brain, hippocampus"  
/clone lib="NIH MGC\_95"  
/lab host="DH10B"  
/note="Vector: pBluescript"  
166. 777  
/codon\_start=1  
/product="Unknown (protein for MGC:87227)"  
/protein id="AAH66929.1"  
/db\_xref="GI:44890461"  
/translation="MAQRKNAKSSGSSGSGSAGSSSPGARRETHKGHGKNG  
RKGLSGTSFTFWVIALLVGVTWVAVVWEDVDVYEVLAKEADFRYNLSEVLQGLK  
GIYDADGDFDQDAKVLGDKGNENIDSLSEVLNLAESSDWFYGLFLSLYD  
IMTFPELMEEESEADTGVDGTSQNEGVGKTCVILDHQNQ"  
319. 558  
/note="Asp-B-Hydro N; Region: Aspartyl beta-hydroxylase  
N-terminal region. This family includes the N-terminal  
regions of the junctin, junctate and aspartyl  
beta-hydroxylase proteins. Junctate is an integral ER/SR  
membrane calcium binding protein, which comes from an  
alternatively spliced form of the same gene that generates  
aspartyl beta-hydroxylase and junctin. Aspartyl  
beta-hydroxylase catalyses the post-translational  
hydroxylation of aspartic acid or asparagine residues  
contained within epidermal growth factor (EGF) domains of  
proteins"  
/db\_xref="CDD:pfam05279"

**ORIGIN**

Query Match 9.2%; Score 213; DB 9; Length 1608;  
Best Local Similarity 99.6%; Pred. No. 5.5e-105;  
Matches 263; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACCGTGCAATGCCAGCGTGAAGTGAATGCCAGAGCGGCGGCAACAGCAGCAGCAGCG 60  
Db 155 CGGACCGTGCAATGCCAGCGTGAAGTGAATGCCAGAGCGGCGGCAACAGCAGCAGCAGCG 214  
Qy 61 GCTCCGCGAGCGGTAGCAGAGTGGCGAGCAGCAGCCCGCGGCGCCCGGAGAGACAA 120  
Db 215 GCTCCGCGAGCGGTAGCAGAGTGGCGAGCAGCAGCCCGCGGCGCCCGGAGAGACAA 274  
Qy 121 AGCATGGAGGACACAAGATGGGGAAGCGGACTCTCGGGAACCTTCATCTTCAGGT 180  
Db 275 AGCATGGAGGACACAAGATGGGGAAGCGGACTCTTCAGGAACCTTCATCTTCAGGT 334

**FEATURES**  
source

181 GGTATTATGTCGATTGTCATTGCTGGCGCTCTGGACATCTGTAGCTGTGTTGGTTGATC 240  
|||||  
335 GGTATTATGTCGATTGTCATTGCTGGCGCTCTGGACATCTGTAGCTGTGTTGGTTGATC 394  
|||||

Qy 241 TTGTTGACTATGAGGAAGTTCTAG 264  
|||||  
Db 395 TTGTTGACTATGAGGAAGTTCTAG 418  
|||||

**RESULT** 19  
AC021323  
LOCUS  
DEFINITION Homo sapiens clone RP11-23M4, WORKING DRAFT SEQUENCE, 2 ordered  
pieces.  
AC021323  
AC021323.5 GI:14030007  
HTG; HTGS PHASE2; HTGS\_DRAFT; HTGS\_FULLTOP.  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 132868)  
Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
Homo sapiens, clone RP11-23M4  
Unpublished  
2 (bases 1 to 132868)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Becker, R., Bieda, F.,  
Boguslavsky, I., Bouckhalter, B., Brown, A., Burket, G., Castle, A.,  
Choepel, Y., Collangelo, M., Collins, S., Collymore, A., Cooke, P.,  
DeArelano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J.,  
Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,  
Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,  
Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K.,  
Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,  
McPheeters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,  
Norman, C. H., O'Connor, T., O'Donnell, P., Olivari, T. M., Peterson, K.,  
Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,  
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,  
Zimmer, A. and Zody, M.

**Direct Submission**  
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On May 13, 2001 this sequence version replaced gi:12229503.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L4398  
Center clone name: 23 M 4  
----- Summary Statistics  
Sequencing vector: M13; M77B15; 3% of reads  
Sequencing vector: Plasmid; n/a; 97% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 132571 bases at least Q40  
Consensus quality: 132698 bases at least Q30  
Consensus quality: 132753 bases at least Q20  
Insert size: 127000; agarose-fp  
Insert size: 132768; sum-of-contigs  
Quality coverage: 20.3 in Q20 bases; agarose-fp  
Quality coverage: 19.5 in Q20 ba.  
\* NOTE: This is a working draft sequence. It currently  
\* consists of 2 contigs. Gaps between the contigs

\* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submittor.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

\* 1 130383: contig of 130383 bp in length  
 \* 130384 130483: gap of 100 bp  
 \* 130484 132868: contig of 2385 bp in length.

#### FEATURES

Location/Qualifiers  
 1. 132868  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="RP11-23M4"  
 /clone\_lib="RPC1-11 Human Male BAC"  
 1. 130383  
 /note="assembly\_fragment"  
 clone\_end:SP6  
 vector\_side:left  
 130484. 132868  
 /note="assembly\_fragment"  
 clone\_end:T7  
 vector\_side:right"

#### misc\_feature

#### misc\_feature

#### ORIGIN

Query Match 8.2%; Score 190; DB 2; Length 132868;  
 Best Local Similarity 100.0%; Pred. No. 2,7e-92;  
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 2135 CAGGACCTGGGAGGAGCAAGTGCTCATCTTTGATGACCTCTTTGAGCAGGATG 2194  
 Db 19703 CAGGACCTGGGAGGAGCAAGTGCTCATCTTTGATGACCTCTTTGAGCAGGATG 19762  
 Qy 2195 GCAGGATCGCTCATCTTTCCGGCTGATTCATCGTGGATGTGTGGCATCCGGAATGAC 2254  
 Db 19763 GCAGGATCGCTCATCTTTCCGGCTGATTCATCGTGGATGTGTGGCATCCGGAATGAC 19822  
 Qy 2255 ACCACAGCAGAGCAGCGACCTTCACGCAATTTAGCATGAATTCATGCAAGCTTGGGAAC 2314  
 Db 19823 ACCACAGCAGAGCAGCGACCTTCACGCAATTTAGCATGAATTCATGCAAGCTTGGGAAC 19882  
 Qy 2315 TCTGGAGAGA 2324  
 Db 19883 TCTGGAGAGA 19892

#### RESULT 20

AC090094  
 LOCUS 158404 bp DNA linear PRI 15-MAR-2002  
 DEFINITION Homo sapiens chromosome 8, clone RP11-252C19, complete sequence.  
 AC090094  
 VERSION AC090094.5 GI:19482342  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 158404)  
 Homo sapiens chromosome 8, clone RP11-252C19

#### REFERENCE

#### AUTHORS

#### JOURNAL

#### AUTHORS

Birren,B., Linton,L., Nussbaum,C., Lander,E., Allen,N., Anderson,S.,  
 Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B., Brown,A.,  
 Camarata,J., Campopiano,A., Choepel,Y., Collangelo,M., Collins,S.,  
 Collumore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,  
 Dodge,S., Fero,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J.,  
 Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,  
 Jones,C., Karatas,A., LaRoque,K., Lamazares,R., Lander,E.,  
 Lehoczy,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,

Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,  
 McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,  
 Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,  
 O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,  
 Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,  
 Ribback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,  
 Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,  
 Souney,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N.,  
 Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Travers,M.,  
 Travis,N., Triggilio,J., Vassiliev,H., Velei,R., Vo,A., Wilson,B.,  
 Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L.,  
 Zember,L., Zimmer,A. and Zody,M.

#### TITLE

#### JOURNAL

#### REFERENCE

#### AUTHORS

Submitted (14-FEB-2001) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 158404)  
 Birren,B., Linton,L., Nussbaum,C., Lander,E., Ali,A., Allen,N.,  
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,  
 Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,  
 Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collumore,A.,  
 Cooke,P., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
 Fero,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,  
 Lander,E., Lehoczy,J., Levine,R., Major,J., Marquis,N., Meneus,L.,  
 Maclean,C., Macdonald,P., McKernan,K., Meldrim,J., Mlenga,V.,  
 Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,  
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,  
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
 Raymond,C., Retta,R., Ribback,M., Riley,R., Rise,C., Rogov,P.,  
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,  
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
 Topham,K., Travers,M., Travis,N., Triggilio,J., Vassiliev,H.,  
 Velei,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (15-MAR-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Mar 15, 2002 this sequence version replaced gi:14626341.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www.seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L12758  
 Center clone name: 252\_C\_19  
 -----

#### TITLE

#### JOURNAL

#### COMMENT

Submitted (15-MAR-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Mar 15, 2002 this sequence version replaced gi:14626341.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www.seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L12758  
 Center clone name: 252\_C\_19  
 -----

#### FEATURES

#### source

Location/Qualifiers  
 1. 158404  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosomes="8"  
 /map="8"  
 /clone="RP11-252C19"  
 /clone\_lib="RPC1-11 Human Male BAC"  
 733. 864  
 /rpt\_family="L2"  
 complement(1037..1417)  
 /rpt\_family="MLT2D"  
 complement(1697..1825)  
 /rpt\_family="LIME"  
 1826..2120  
 /rpt\_family="AluX"  
 1960..1969  
 /note="x30 qual SNGL region"

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region

#### repeat\_region



JOURNAL Submitted (13-JAN-2000) Life Science, Kwangju Institute of Science and Technology, Kwangju 500-712, Korea

FEATURES  
source Location/Qualifiers  
1..855  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="heart"  
112..744  
/note="junctional sarcoplasmic reticulum protein"  
/codon\_start=1  
/product="junction"  
/protein\_id="AAF82246.1"  
/db\_xref="GI:8979460"  
/translation="MAEDKETKHGHKNGKGLSGTSTFTFMVIALIGVWTSVAVV  
WFDLVYEVILGKGIYDADGDGDFDIDAKVLLRSPSVAKRKTAKVKELTKELK  
KEKEKPEKSKESKNEERKKGKEDVEKDKKIADLSPKSPKGGKDKREKVDLEKS  
AKTKENRKKSTNMKDVSSKWSKDKDRKESRSSTRYAHLTKGNTQKRG"

CDS  
7.4%; Score 172; DB 9; Length 855;  
Query Match 100.0%; Pred.No. 2.2e-82;  
Best Local Similarity 0; Mismatches 0; Gaps 0;  
Matches 172; Conservative 0; Indels 0;

ORIGIN  
162 GGAACATTCATCTTCACGTGTTATGATGATTCATTCCTGGCGCTGGACATCTGTA 221  
Dbb 175 GGAACATTCATCTTCACGTGTTATGATGATTCATTCCTGGCGCTGGACATCTGTA 234  
Qy 222 GCTGCTGTTGTTGATCTGTTGATCTGATGAGGAGTTCTAGGAAACTAGGAATCTAT 281  
Dbb 235 GCTGCTGTTGTTGATCTGTTGATCTGATGAGGAGTTCTAGGAAACTAGGAATCTAT 294  
Qy 282 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 333  
Dbb 295 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 346

RESULT 22  
AC067881/c  
LOCUS  
DEFINITION Homo sapiens chromosome 8, clone RP11-687K23, complete sequence.  
AC067881  
VERSION AC067881.5 GI:15799630  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 186676)  
Birren.B., Linton.L., Nusbaum.C., Lander.E., Abraham.H., Allen.N.,  
Anderson.S., Baldwin.J., Barna.N., Bastien.V., Bada.F.,  
Boguslavsky.L., Bouckhgalter.B., Brown.A., Burkett.G.,  
Campopiano.A., Castle.A., Choepel.Y., Colangelo.M., Collins.S.,  
Collumore.A., Cooke.P., DeArellano.K., Dewar.K., Diaz.J.S.,  
Dodge.S., Domino.M., Doyle.M., Ferreira.P., FitzHugh.W., Gage.D.,  
Galagan.J., Gardyna.S., Ginde.S., Goyette.M., Graham.L.,  
Grand-pierre.N., Grant.G., Hagos.B., Heaford.A., Horton.L.,  
Howland.J.C., Iliev.I., Johnson.R., Jones.C., Kann.L., Karatas.A.,  
Klein.G., LaRoque.K., Lamazares.R., Landers.T., Lehoczyk.J.,  
Levine.R., Lieu.C., Liu.G., Locke.K., Macdonald.P., Marquis.N.,  
McCarthy.M., McSwan.P., McGurk.A., McKernan.K., McPheeters.R.,  
Meldrin.J., Meneus.L., Mihova.T., Miranda.C., Mlenga.V., Morrow.J.,  
Murphy.T., Naylor.J., Norman.C.H., O'Connor.T., O'Donnell.P.,  
O'Neill.D., Oliver.T.M., Oliver.J., Peterson.K., Pierre.N.,  
Pisani.C., Pollara.V., Raymond.C., Riley.R., Rogov.P., Rothman.D.,  
Roy.A., Santos.R., Schauer.S., Severy.P., Spencer.B.,  
Stange-Thomann.N., Stojanovic.N., Subramanian.A., Talamas.J.,  
Testfaye.S., Theodore.J., Tirrell.A., Travers.M., Trigilio.J.,  
Vassiliev.H., Viel.R., Vo.A., Wilson.B., Wu.X., Wyman.D., Ye.W.J.,

Young.G., Zairoun.J., Zimmer.A. and Zody.M.  
Direct Submission  
Submitted (27-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 186676)  
Birren.B., Linton.L., Nusbaum.C., Lander.E., Ali.A., Allen.N.,  
Anderson.S., Barna.N., Bastien.V., Boguslavsky.L., Bouckhgalter.B.,  
Brown.A., Camarata.J., Campopiano.A., Chang.J., Chazaro.B.,  
Choepel.Y., Colangelo.M., Collins.S., Collumore.A., Cooke.A.,  
Cooke.P., DeArellano.K., Dewar.K., Diaz.J.S., Dodge.S., Faro.S.,  
Ferreira.P., FitzHugh.W., Gage.D., Galagan.J., Gardyna.S.,  
Ginde.S., Goid.S., Goyette.M., Graham.L., Grand-pierre.N.,  
Hagos.B., Heaford.A., Horton.L., Hulme.W., Iliev.I., Johnson.R.,  
Jones.C., Kanat.A., Karatas.A., Kells.C., LaRoque.K.,  
Lamazares.R., Landers.T., Lehoczyk.J., Levine.R., Liu.G.,  
MacLean.C., Macdonald.P., Majox.J., Marquis.N., Matthews.C.,  
McCarthy.M., McSwan.P., McKernan.K., McPheeters.R., Meldrin.J.,  
Meneus.L., Mihova.T., Mlenga.V., Murphy.T., Naylor.J., Nguyen.C.,  
Norbu.C., Norman.C.H., O'Connor.T., O'Donnell.P., O'Neill.D.,  
Oliver.J., Peterson.K., Phunkhang.P., Pierre.N., Pollara.V.,  
Raymond.C., Retta.R., Rieback.M., Riley.R., Rise.C., Rogov.P.,  
Roman.J., Rosetti.M., Roy.A., Santos.R., Schauer.S., Schupback.R.,  
Seaman.S., Severy.P., Spencer.B., Stange-Thomann.N., Stojanovic.N.,  
Straham.K., Subramanian.A., Talamas.J., Testfaye.S., Theodore.J.,  
Topham.K., Travers.M., Travis.N., Trigilio.J., Vassiliev.H.,  
Viel.R., Vo.A., Wilson.B., Wu.X., Wyman.D., Ye.W.J., Young.G.,  
Zairoun.J., Zembek.L., Zimmer.A. and Zody.M.  
Direct Submission  
Submitted (28-SEP-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Sep 28, 2001 this sequence version replaced gi:15422051.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L8165  
Center clone name: 687\_K\_23  
----- Location/Qualifiers  
1..186676  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"  
/clone="RP11-687K23"  
/clone\_lib="RP11-11 Human Male BAC"  
243..278  
/rpt\_family="(TG)n"  
complement(1919..2154)  
/rpt\_family="MER96B"  
complement(2155..2456)  
/rpt\_family="Alusx"  
3358..4039  
/rpt\_family="Tigger1"  
4046..4839  
/rpt\_family="Tigger1"  
complement(5527..7135)  
/rpt\_family="L2"  
complement(7894..8215)  
/rpt\_family="MER47A"  
9173..9513  
/rpt\_family="MER2"  
9914..9946  
/rpt\_family="AT-rich"  
10203..10660  
/rpt\_family="L2"  
10661..12746



```

OY      487  TACACGACGACATG 501
Db      121  TACACGACGACATG 135

RESULT 24
LOCUS   CR626894/c
DEFINITION Human DNA sequence from clone XX-NCIH2171_3M21, complete sequence.
ACCESSION CR626894
VERSION   CR626894.1
KEYWORDS  HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens

REFERENCE
1 (bases 1 to 126392)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Dunn, M.
Direct Submission
Submitted (27-JUL-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CH10 1SA, UK. E-mail enquiries:
clonerequest@sanger.ac.uk
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
XX-NCIH2171_3M21
is from a Cancer cell lines BACS library VECTOR: pBACe3.6_BanHI.

FEATURES             source
1..126392
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /clone="XX-NCIH2171_3M21"
    /clone_lib="Cancer Cell lines BACS"

ORIGIN
Query Match      5.6%; Score 130; DB 9; Length 126392;
Best Local Similarity 100.0%; Pred. No. 3.2e-59;
Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      501  GTTGGGAGAGACTTGGCAACGAGAGATGACCCACGAGAGACCAACGAGAGAT 560
Db      42963 GTTGGGAGAGAGACTTGGCAACGAGAGATGACCCACGAGAGACCAACGAGAGAT 42904

OY      561  GATGAGTTTCTTATGGCGACTGATGATGATGATGATGATGATGATGATGATGATGAT 620
Db      42903 GATGAGTTTCTTATGGCGACTGATGATGATGATGATGATGATGATGATGATGATGAT 42844

OY      621  TCTCATGAAG 630
Db      42843 TCTCATGAAG 42834

```

```

RESULT 25
AC091173
LOCUS   AC091173
DEFINITION Homo sapiens chromosome 8, clone RP11-280G9, linear PRI 16-OCT-2002
AC091173
ACCESSION AC091173
VERSION   AC091173.10
KEYWORDS  HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens

REFERENCE
1 (bases 1 to 142265)
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A.,
Camarata, J., Campopiano, A., Chang, J., Choquel, Y., Collangelo, M.,
Collins, S., Collymore, A., Cooke, P., DeArellano, K., Dewar, K.,
Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W.,
Iliev, I., Johnson, R., Jones, C., Karatas, A., LaRoque, K.,
Lamaze, R., Landers, T., Lechoczky, J., Levine, R., Liu, G.,
MacLean, C., Macdonald, P., Marquis, N., Matthews, C., McCarthy, M.,
McEwan, P., McKernan, K., McPheeters, R., Meldrum, J., Meneus, L.,
Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S.,
Severy, P., Sougnez, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S.,
Theodore, J., Travers, N., Travis, N., Triglio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (03-APR-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 142265)
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choquel, Y., Collymore, A.,
Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mienga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (08-SEP-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 142265)
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,
Camarata, J., Chang, J., Chazaro, B., Choquel, Y., Collymore, A.,
Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mienga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (08-SEP-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 142265)

```







```
/codon_start=1
/product="aspartyl (asparaginyl) beta hydroxylase"
/protein_id="AAA03563.1"
/db_xref="GI:162694"
/translation="MAPRKNAGGGGSSSSSGPTGCTGGSSSPGARRETKQGL
KNGKGGSGSFTFWVIALLGWTSVAWVDFLDVYEVILAKADFRYNLSVLQ
KGLGIDAGDGDVDDAALLGLKEKPAKPTVPPEADMPWLEDDQVLESGRQ
IEDEVYQVQSLDVTYSEGENLPQPEGAPELQPDHVFVGSDDADRYEPMTGA
VHESTESYHLEETASPAYSDQMEDMYEQENPSSSEPVVDDAERYVQETDVTVD
YDEQHAVDNSNTILEPHMPPEAEQEQVPEITNKADPEKGGKVKKKPKLLNKFD
KTIKEDLAARKEKRIEAVNAFELVRKYPOSPGARYGAQCEDDLAEKRRSNE
ILRAIETVQAASLPDPTDLVLSLRSDRQDFLGHMRGSLTLTKLVQLFPDPT
ALKNDLGVYLLIGDPSAKVYEVLSVTPNDGFAKHYGFILKAQNKIAETPYLK
EGIESGPDGTGRFVHLGDAMORVGNKEARYVVELGHORGHGFPASVWQSLYVQGL
KAQPWTPKETGYELVKSLEFNWKLJDEGLAMDRTHGLFLPEDENLEKGDWSOF
TLWQQRKNACGAPKTSLLDKPFETTCRRGQIKYSIMHFGTHVPHWTGTNCR
LRMHLGLVPEKGGKRICANETRWEEGKVLIFDSDSEHEVWQDAASFLRIFVDVWH
PELTPHRRSLPAI"
```

## ORIGIN

Query Match 2.7%; Score 62; DB 4; Length 2739;

Best Local Similarity 100.0%; Pred. No. 9.5e-22; Mismatches 0; Indels 0; Gaps 0;

Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 273 GGAATCTATGATGCTGATGGATGGAGATTTTGTGATGATGCGCAAAAGTTTATTA 332  
|||||  
Db 504 GGAATCTATGATGCTGATGGATGGAGATTTTGTGATGATGCGCAAAAGTTTATTA 563  
|||||

QY 333 GG 334

||

Db 564 GG 565

## RESULT 29

CQ541435  
LOCUS CQ541435 60 bp DNA linear PAT 30-JAN-2004  
DEFINITION Sequence 11070 from Patent WO0210449.

ACCESSION CQ541435

VERSION CQ541435.1 GI:41507699

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 Shoshan, A., Wasserman, A., Mintz, E., Mintz, L. and Faigler, S.  
Oligonucleotide library for detecting rna transcripts and splice  
variants that populate a transcriptome

JOURNAL Patent: WO 0210449-A 11070 07-FEB-2002;  
CompuGen Inc. (US)

FEATURES

Location/Qualifiers

source

1..60

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

## ORIGIN

Query Match 2.6%; Score 60; DB 6; Length 60;

Best Local Similarity 100.0%; Pred. No. 1.2e-20; Mismatches 0; Indels 0; Gaps 0;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 672 GACTGTAAATCAGGATGGAAGATGATGTTCTGAGCAGGAAATCCAGATTCCAGTGAA 731  
|||||  
Db 1 GACTGTAAATCAGGATGGAAGATGATGTTCTGAGCAGGAAATCCAGATTCCAGTGAA 60  
|||||

## RESULT 30

AX771609/c  
LOCUS AX771609 2208 bp DNA linear PAT 02-JUL-2003  
DEFINITION Sequence 322 from Patent WO03004646.

ACCESSION AX771609

VERSION AX771609.1 GI:32438407

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

O'Mahony, D.J., Byrne, D., Brayden, D., Lambkin, I. and Higgins, L.  
Genetic analysis of Peyer's patches and M cells and methods and  
compositions targeting Peyer's patches and M cell receptors

JOURNAL

Patent: WO 03004646-A 322 16-JAN-2003;  
ELAN CORPORATION, PLC (IE)

FEATURES

Location/Qualifiers

source

1..2208

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

## ORIGIN

Query Match 2.5%; Score 59; DB 6; Length 2208;

Best Local Similarity 100.0%; Pred. No. 4.3e-20; Mismatches 0; Indels 0; Gaps 0;

Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2239 GGCATCCGGAACACACACAGAGAGAGCGAGCCTTCCAGCAATTAGCATGAATTC 2297  
|||||  
Db 59 GGCATCCGGAACACACACAGAGAGAGCGAGCCTTCCAGCAATTAGCATGAATTC 1  
|||||

## RESULT 31

HSATPBR/c

LOCUS

DEFINITION Human mRNA for Na/K-ATPase beta subunit.

ACCESSION X03747

VERSION X03747.1 GI:28932

KEYWORDS ATPase; glycoprotein.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

TITLE

1 (bases 1 to 2208)  
Kawakami, K., Nojima, H., Ohta, T. and Nagano, K.  
Molecular cloning and sequence analysis of human Na,K-ATPase  
beta-subunit

JOURNAL

Nucleic Acids Res. 14 (7), 2833-2844 (1986)

MEDLINE

PUBMED

COMMENT Data kindly reviewed (10-AUG-1986) by K. Kawakami.

FEATURES

Location/Qualifiers

source

1..2208

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

CDS

127..1038

/note="unnamed protein product; put. Na/K-ATPase beta (aa

1-303)"

/codon\_start=1

/protein\_id="CAA27385.1"

/db\_xref="GI:28933"

/db\_xref="GOA:P05026"

/translation="MARGKAKEGSWKFIWNSEKFLGRTGGSWFKILLFYVIFYG

CLAGFTIGTQVMLLTISEFKPTQDRVAPPLGTFQIQKTEISFRPNDPKSYEAYV  
LNIYRLEKYKDSARDMDIPEDCGDVSPKRGEDFNHERGERKVRFKLEWLGNC  
GLNDETGYKGGKPCIIKLNRVLGKPKPKPKNESLETYPVMKYNPNVLPVQCTGKRD  
EDDKVGNVEYFGLGNSPGFPLOYVYVYKLLQPKYLQPLAVQFTNLTMTDIRIEC  
KAYGENIGYSEKDRFQGRFDFVKIEVKS"

misc\_feature

598..600

/note="pot. glycosylation site"

misc\_feature

703..705

/note="pot. glycosylation site"

misc\_feature

919..921

/note="pot. glycosylation site"

misc\_feature

1485..1490

/note="pot. polyA signal"

misc\_feature

2001..2006

```

misc_feature      /note="pot. polyA signal"
2026..2031       /note="pot. polyA signal"
2187..2193       /note="pot. polyA signal"
ORIGIN
Query Match      2.5%; Score 59; DB 9; Length 2208;
Best Local Similarity 100.0%; Pred. No. 4.3e-20;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2239 GGCATCCGGAACATGACACACAGAGAGAGCGAGCGCTTCCAGCAATTTAGCATGAATTC 2297
DB 59 GGCATCCGGAACATGACACACAGAGAGAGCGAGCGCTTCCAGCAATTTAGCATGAATTC 1

RESULT 32
AC143562
LOCUS      Macaca mulatta clone CH250-267C14, DNA linear HTG 09-APR-2003
DEFINITION Macaca mulatta clone CH250-267C14, *** SEQUENCING IN PROGRESS ***.
AC143562
VERSION    HTG; HTGS_PHASE2; HTGS_PGI.
KEYWORDS   HTG; HTGS_PHASE2; HTGS_PGI.
SOURCE     Macaca mulatta (rhesus monkey)
ORGANISM   Macaca mulatta

REFERENCE
AUTHORS    Ceurots M. and Milosavljevic A.
TITLE      Pooled genomic indexing (PGI): mathematical analysis and experiment
           design (in) Guigo, R. and Gusfield, D. (Eds.);
           ALGORITHMS IN BIOINFORMATICS, SECOND INTERNATIONAL WORKSHOP, WABI
           2002, ROME, ITALY, SEPTEMBER 17-21, 2002, PROCEEDINGS: 10-28;
           Springer (2002)
           2 (bases 1 to 177614)
MILosavljevic A., Sodergren, E., Ceurots, M., Li, B., Jackson, A.R.,
Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C., Alsbrooks, S.L.,
Amaratunge, H.C., Are, J.R., Ayale, M., Banks, T., Barbara, J.,
Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J.,
Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C.,
Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,
Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G.F., Chen, R.,
Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C., Cleveland, C.D.,
Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C.,
Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,
Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H.,
Dugan-Rocha, S., Durbin, K.J., Egan, A., Earnhart, C., Edwards, C.C.,
Elhaj, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D.,
Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J.,
Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W.,
Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K.,
Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O.,
Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homs, F.,
Howard, S., Huber, J., Hulyk, S., Hume, J., Ioshikhes, I., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karleson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L.C.,
Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,
Loulsegh, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
Ma, J., Maheshwari, M., Mapua, P., Marondel, I., Martin, R.,
Martindale, A., Martinez, S., Massey, E., Mawhinney, E., McLeod, M.P.,
Meador, M., Mei, G., Merscher, S., Metzger, M., Miller, A., Miner, G.,
Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K.T., Morgan, M.,
Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N.,
Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenwo, S.,
Ogulu, M., Okwono, G., Oragunye, N., Owiedo, R., Pace, A., Payton, B.,
Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L.,
Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M.,
Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shim, C.,
Shooshtari, N., Sison, I., Sodergren, E., Sonaike, T., Sparks, A.,
Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A.,

```

Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williams, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y. F., Zhou, J., Zorrilla, S., Kuchelapatti, R., Weinstein, G. and Gibbs, R.

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## JOURNAL

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

## COMMENT

```

/db_xref="GI:11878110"
/translation="MAPRKNAKGGGNGSSSGSGSGSGSPSTGTSSGSSSSPGARRE
AKHGCHKNGRRGGISGGSFFTWFMVIALLGVTSAVVWFDLVDYEEVLKGLGVYDAD

```

```
GGGDFDVAADKAVLLGLKERSPSERTPPPEAEETHAELECAPEGADIONVEDEVKEQIO
SLIQSVHTDHLADGLAGEPOEVEDFLTVTDSDDRFEDLEPGTVHERIEDTYHVE
DTASQNHNDMEEMTNEQNSDESEAVTDAGVLLPHAEVRHQDYDPVPESEHGV
RKSNDNTIDDSIIIEEINVASVEEQDTPPVKKKKPKLLNKFDKTIKAEIDAAELRK
RGIKEEAVNAPEELVRKYPOSPRARYGKACQEDDLAEQRNSVRLAIETYGEADL
DPADTDLVKLSKRSRQOFLGHMRGSLTLQRLVOLFPDSTTLKNDLGVGILLGD
NDSAKVVEELVNTVPDNGFAKHVYGFILKAQNKISIPFIKSGIESGDPDGTGFE
VPHLGDAMORVGNKEAYKVELGHKGFASVQRLSNVNLKAQPMWTPRETGGT
LVKSLERWKLIRDEGLVMDKAGLFLPEDENLRKGDWSQFTLWQOGRKNENACKG
AKTICALLKESITGCRGQIKYSIMHFGTHVPHGTGTCNRLMRHLGLVPIKEGCK
ICANETITWEEGKVLIFDDSFEEHVQDASSFRILFIVDMVHPELTPOQRSLPAI"
2395..4419
/gene="Asph"
4379..4384
/gene="Asph"

3'UTR
polyA_signal
ORIGIN
Query Match
Best Local Similarity 100.0%; Pred. No. 7.3e-07; Length 4419;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2208 TCCTTCGGCTGATATTCATCGTGGATGTGGCA 2242
|||||
Db 2314 TCCTTCGGCTGATATTCATCGTGGATGTGGCA 2348

RESULT 36
H010058S11 6024 bp DNA linear ROD 13-MAR-2001
LOCUS Mus musculus aspartyl beta-hydroxylase (ASPH) gene, exon 24 and
DEFINITION complete cds.
ACCESSION AF289215
VERSION AF289215.2 GI:13310928
KEYWORDS
SOURCE 11 of 11
MUS musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Dinchuk, J.E., Henderson, N.L., Burn, T.C., Huber, R., Ho, S.P.,
Link, J., O'Neill, K.T., Focht, R.J., Scully, M.S., Hollis, J.M.,
Hollis, G.F. and Friedman, P.A.
Aspartyl beta-hydroxylase (Asph) and an evolutionarily conserved
isoform of Asph missing the catalytic domain share exons with
junction
J. Biol. Chem. 275 (50), 39543-39554 (2000)
MEDLINE 20564328
PUBMED 10956665
REFERENCE 2 (bases 1 to 6024)
Henderson, N.L., Dinchuk, J.E., Burn, T.C., Hollis, G.F. and
Friedman, P.A.
Direct Submission
Submitted (25-JUL-2000) Applied Biotechnology, DuPont
Pharmaceuticals Company, PO BOX 80400, Wilmington, DE 19880, USA
3 (bases 1 to 6024)
Henderson, N.L., Dinchuk, J.E., Burn, T.C., Hollis, G.F. and
Friedman, P.A.
Direct Submission
Submitted (13-MAR-2001) Applied Biotechnology, DuPont
Pharmaceuticals Company, PO BOX 80400, Wilmington, DE 19880, USA
REMARK Sequence update by submitter
COMMENT On Mar 13, 2001 this sequence version replaced gi:11692644.
FEATURES
source
1..6024
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
order(AF289205:1:1772...2483,AF289206:1:1..1951,
AF289207:2:1..31245,AF289208:1:1..2633,AF289209:1:1..2013,
AF289210:1:1..5161,AF289211:1:1..6629,AF289212:1:1..2335,
AF289213:1:1..5657,AF289214:1:1..12095,1..>6000)
/gene="ASPH"

mRNA
join(AF289205:1:1772...2075,AF289206:1:1105...1173,
AF289207:2:1425...4514,AF289207:2:4438...4512,
AF289207:2:7694...7807,AF289207:2:9820...9852,
AF289207:2:10948...11004,AF289207:2:12520...12567,
AF289207:2:12997...13029,AF289207:2:15883...15921,
AF289207:2:16600...16635,AF289207:2:20179...20223,
AF289208:1:1520...1605,AF289209:1:292...378,
AF289210:1:2354...2504,AF289210:1:4279...4415,
AF289211:1:1764...1862,AF289211:1:4834...4923,
AF289212:1:428...565,AF289213:1:2049...2184,
AF289214:1:5360...5451,AF289214:1:5805...5938,1665...>6000)
/gene="ASPH"
/product="aspartyl beta-hydroxylase"
join(AF289205:1:1946...2075,AF289206:1:1105...1173,
AF289207:2:1425...4514,AF289207:2:4438...4512,
AF289207:2:7694...7807,AF289207:2:9820...9852,
AF289207:2:10948...11004,AF289207:2:12520...12567,
AF289207:2:12997...13029,AF289207:2:15883...15921,
AF289207:2:16600...16635,AF289207:2:20179...20223,
AF289208:1:1520...1605,AF289209:1:292...378,
AF289210:1:2354...2504,AF289210:1:4279...4415,
AF289211:1:1764...1862,AF289211:1:4834...4923,
AF289212:1:428...565,AF289213:1:2049...2184,
AF289214:1:5360...5451,AF289214:1:5805...5938,1665...1815)
/gene="ASPH"
/codon_start=1
/product="aspartyl beta-hydroxylase"
/protein_id="AAG39913.1"
/db_xref="GI:11692645"
/translation="MAPRKNAGCGGSSSSGSGSGSGSPSTGSGSSSSSPFGARRG
KLVYDADGDCDFDVAADKAVLLGLKERSPSERTPPPEAEETHAELECAPEGADIONVE
DEVKEQIQSLIQSVHTDHLADGLAGEPOEVEDFLTVTDSDDRFEDLEPGTVHERIE
DITYHVEDTASQNHNDMEEMTNEQNSDESEAVTDAGVLLPHAEVRHQDYDPVPESE
HGVGRKSGEISNTIDDSIIIEEINVASVEEQDTPPVKKKKPKLLNKFDKTIKAEI
DAAEKLRKGIKEEAVNAPEELVRKYPOSPRARYGKACQEDDLAEQRNSVRLAIET
YGEADLDPADTDLVKLSKRSRQOFLGHMRGSLTLQRLVOLFPDSTTLKNDLGVGIL
LDNDSAKVVEELVNTVPDNGFAKHVYGFILKAQNKISIPFIKSGIESGDPDGTGFE
VPHLGDAMORVGNKEAYKVELGHKGFASVQRLSNVNLKAQPMWTPRETGGTGVYV
PRETGVYVLSLERNWKLIRDEGLVMDKAGLFLPEDENLRKGDWSQFTLWQOGRKN
ENACKGAPKTCALLKESITGCRGQIKYSIMHFGTHVPHGTGTCNRLMRHLGL
VPIKEGCKIRCANETRTWEEGKVLIFDDSFEEHVQDASSFRILFIVDMVHPELTPO
QRRSLPAI"
1665..1815
/gene="ASPH"
number=24
3801..3806
/gene="ASPH"
5995..6000
/gene="ASPH"

exon
polyA_signal
polyA_signal
ORIGIN
Query Match
Best Local Similarity 100.0%; Pred. No. 7.3e-07; Length 6024;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2208 TCCTTCGGCTGATATTCATCGTGGATGTGGCA 2242
|||||
Db 1735 TCCTTCGGCTGATATTCATCGTGGATGTGGCA 1769

RESULT 37
AF289487
LOCUS Mus musculus
DEFINITION Mus musculus aspartyl beta-hydroxylase 6.6 kb transcript (Asph)
mRNA, complete cds; alternatively spliced.
ACCESSION AF289487
VERSION AF289487.1 GI:11878111
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 6629)
REFERENCE
```







```

ORIGIN
  /db_xref="CDD:pfam05279"

Query Match      1.4%; Score 32; DB 10; Length 1763;
Best Local Similarity 100.0%; Pred. No. 3.2e-05;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 303 TTTGATGTGGATGATGCCAAAGTTTATTAGG 334
Db 562 TTTGATGTGGATGATGCCAAAGTTTATTAGG 593

RESULT 41
H010058S02      1951 bp      DNA      linear      ROD 13-MAR-2001
LOCUS      Mus musculus aspartyl beta-hydroxylase (ASPH) gene, exon 3.
DEFINITION      AF289206
ACCESSION      AF289206.1 GI:11692635
VERSION
KEYWORDS
SEGMENT
SOURCE      Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS      Dinchuk, J.E., Henderson, N.L., Burn, T.C., Huber, R., Ho, S.P.,
              Link, J., O'Neill, K.I., Focht, R.J., Scully, M.S., Hollis, J.M.,
              Hollis, G.F. and Friedman, P.A.
              Aspartyl beta-hydroxylase (Asph) and an evolutionarily conserved
              isoform of Asph missing the catalytic domain share exons with
              junctin
              J. Biol. Chem. 275 (50), 39543-39554 (2000)
JOURNAL
MEDLINE      20564328
PUBMED
REFERENCE      2 (bases 1 to 1951)
AUTHORS      Henderson, N.L., Dinchuk, J.E., Burn, T.C., Hollis, G.F. and
              Friedman, P.A.
              Direct Submission
              Submitted (25-JUL-2000) Applied Biotechnology, DuPont
              Pharmaceuticals Company, PO BOX 80400, Wilmington, DE 19880, USA
FEATURES
  source
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10090"
    1105..1173
    /gene="ASPH"
    /number=3
  exon

ORIGIN
  Query Match      1.4%; Score 32; DB 10; Length 1951;
  Best Local Similarity 100.0%; Pred. No. 3.2e-05;
  Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 303 TTTGATGTGGATGATGCCAAAGTTTATTAGG 334
Db 1143 TTTGATGTGGATGATGCCAAAGTTTATTAGG 1174

RESULT 42
AF302653
LOCUS      Mus musculus cardiac junctate 1 mRNA, complete cds.
DEFINITION      AF302653
ACCESSION      AF302653.2 GI:16507230
VERSION
KEYWORDS
SOURCE      Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS      Hong, C.-S. and Kim, D.H.
              Cloning of mouse junctin homologs
TITLE

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
  
```

---

```

JOURNAL
REFERENCE      2 (bases 1 to 2617)
AUTHORS      Hong, C.-S. and Kim, D.H.
TITLE      Direct Submission
JOURNAL
COMMENT      Submitted (05-SEP-2000) Life Science, Kwangju Institute of Science
              and Technology, Puk-gu, Oryong-dong, Kwangju 500-712, South Korea
              On Oct 29, 2001 this sequence version replaced gi:15824389.
FEATURES
  Location/Qualifiers
    1..2617
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="BALB/c"
    /db_xref="taxon:10090"
    67..879
    /codon_start=1
    /product="cardiac junctate 1"
    /protein_id="AA09319.2"
    /db_xref="GI:16507231"
    /translation="MAEDKEAKHGHNRRGIGSGSFTHWVMTVALLGVWTSVAVV
              WFDLVDEEVLKLVYDADGDGDFVDDAKVLLGLKRSFSPRTFFPFAETHAELEE
              QAPEGADIQNVDEDEKVEQIQSLIQESVHTDHLADGLAGEPQVEFDLTVTSDDR
              FEDLEPTGTHEEITDTHVEDTASQNHNDMEEMTNEQNSDPSZAVTADGVLPHAE
              EVRHQDYDEPVYEPSHEGVEISDNTIDSSIISEBINVASVEEQQDTPTDT"
ORIGIN
  Query Match      1.4%; Score 32; DB 10; Length 2617;
  Best Local Similarity 100.0%; Pred. No. 3.3e-05;
  Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 303 TTTGATGTGGATGATGCCAAAGTTTATTAGG 334
Db 271 TTTGATGTGGATGATGCCAAAGTTTATTAGG 302

RESULT 43
BC015281
LOCUS      Mus musculus aspartate-beta-hydroxylase, mRNA (cDNA clone
              IMAGE:4235968), complete cds.
ACCESSION      BC015281
VERSION      BC015281.1 GI:15929715
KEYWORDS
SOURCE      Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS      Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
              Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
              Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
              Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, F.,
              Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
              Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
              Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
              Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
              Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
              McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
              Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
              Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
              Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
              Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
              Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
              Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
              Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E.,
              Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
              Generation and initial analysis of more than 15,000 full-length
              human and mouse cDNA sequences
              Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
              22388257
              MEDLINE
              12477932
              PUBMED
              2 (bases 1 to 2770)
              REFERENCE
              STRAUSBERG, R.
              DIRECT SUBMISSION
              TITLE
  
```



REFERENCE 1 (bases 1 to 198991)  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
TITLE Lovell, J.  
AUTHORS Direct Submission  
JOURNAL Submitted (01-OCT-2002) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
COMMENT On Oct 2, 2002 this sequence version replaced gi:23395448.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk  
-----

During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: Em, EMBL; Sw,  
SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP  
database can be found at  
http://www.sanger.ac.uk/projects/C\_elegans/wormpep RP23-70D2 is  
from the Rpci-23 Mouse PAC library  
constructed by the group of Pieter de Jong.  
For further details see http://www.chori.org/bacpac/home.htm  
VECTOR: pBACE3.6.

FEATURES  
source  
1..198991  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="4"  
/clone\_lib="RPCI-23"  
/clone\_lib="RPCI-23"

ORIGIN  
Query Match 1.4%; Score 32; DB 10; Length 198991;  
Best Local Similarity 100.0%; Pred. No. 3.4e-05;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 303 TTGTGATGCGATGATGCCAAAGTTTATTAGG 334  
|||||  
DB 20973 TTGTGATGCGATGATGCCAAAGTTTATTAGG 20942  
|||||

RESULT 46  
AF289492 566 bp mRNA linear ROD 18-DEC-2000  
LOCUS Mus musculus asparaginyl beta-hydroxylase cardiac isoform 2 mRNA,  
partial cds; alternatively spliced.  
ACCESSION AF289492  
VERSION AF289492.1 GI:11878121  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 566)  
AUTHORS Dinchuk, J.E., Henderson, N.L., Burn, T.C., Huber, R., Ho, S.P.,  
Link, J., O'Neill, K.T., Focht, R.J., Scully, M.S., Hollis, J.M.,  
Hollis, G.F., and Friedman, P.A.  
TITLE Aspartyl beta-hydroxylase (Asph) and an evolutionarily conserved  
isoform of Asph missing the catalytic domain share exons with

JUNCTION  
J. Biol. Chem. 275 (50), 39543-39554 (2000)  
MEDLINE 20564328  
PUBMED 10956665  
REFERENCE 2 (bases 1 to 566)  
AUTHORS Dinchuk, J.E., Henderson, N.L., Burn, T.C., Hollis, G.F. and  
Friedman, P.A.  
TITLE Direct Submission  
JOURNAL Submitted (26-JUL-2000) Applied Biotechnology, DuPont  
Pharmaceuticals, P.O. Box 80400, Wilmington, DE 19880-0400, USA  
FEATURES  
source  
1..566  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="BALB/c"  
/db\_xref="taxon:10090"  
/chromosome="8"  
/map="8q12 at STS WT-11767"  
/tissue\_type="heart"  
1..124  
125..566  
/note="BAH; asparaginyl beta-hydroxylase; Asph;  
alternatively spliced"  
/codon\_start=1  
/product="aspartyl beta-hydroxylase cardiac isoform 2"  
/protein\_id="AAC40814.1"  
/db\_xref="GI:11878122"  
/translation="MAEDKEAKHGKNGRGISGGSFTTFMVTALGLGWTSTAVV  
WPLDVEYVLGKLVGDGDFDADKVLGEGGLAKTKAKGLKERSPSER  
TFPEAAHAELBEQAEQADIQNVEDEKVEQISLLQESVHTDH"

ORIGIN  
Query Match 1.3%; Score 31; DB 10; Length 566;  
Best Local Similarity 100.0%; Pred. No. 0.00011;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 303 TTGTGATGCGATGATGCCAAAGTTTATTAG 333  
|||||  
DB 329 TTGTGATGCGATGATGCCAAAGTTTATTAG 359  
|||||

RESULT 47  
AF221854 919 bp mRNA linear ROD 02-FEB-2001  
LOCUS Mus musculus cardiac junctin mRNA, complete cds.  
ACCESSION AF221854  
VERSION AF221854.1 GI:12655824  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 919)  
AUTHORS Jones, L.R., Zhang, L., Sanborn, K., Jorgensen, A.O. and Kelley, J.  
TITLE of the 26-kDa calsequestrin binding protein (junctin) from cardiac  
junctional sarcoplasmic reticulum

J. Biol. Chem. 270 (51), 30787-30796 (1995)  
MEDLINE 96107245  
PUBMED 8530521  
REFERENCE 2 (bases 1 to 919)  
AUTHORS Zhang, L. and Jones, L.R.  
TITLE Direct Submission  
JOURNAL Submitted (06-JAN-2000) 47, Sigma Chemical Co., N-3, 3300 South  
Second Street, St. Louis, MO 63118, USA  
FEATURES  
source  
1..919  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/tissue\_type="cardiac muscle"  
143..766  
/note="calsequestrin binding protein"

ORIGIN

Query Match 1.3%; Score 31; DB 10; Length 919;  
 Best Local Similarity 100.0%; Pred. No. 0.00011;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 303 TTTGATGTGGATGATGCCAAAGTTTATTAG 333  
 Db 347 TTTGATGTGGATGATGCCAAAGTTTATTAG 377

RESULT 48

AF223414  
 LOCUS Mus musculus junctin-2 mRNA, complete cds. ROD 19-DEC-2002  
 DEFINITION Mus musculus junctin-2 mRNA, complete cds.  
 ACCESSION AF223414  
 VERSION AF223414.1 GI:27261147  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 988)  
 AUTHORS Hong, C.-S. and Kim, D.H.  
 TITLE Mouse junctin-2 mRNA  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 988)  
 AUTHORS Hong, C. and Kim, D.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-JAN-2000) Life Science, Kwangju Institute of Science and Technology (K-JIST), Puk-gu, Oryong-dong, Kwangju 500-712, South Korea

FEATURES

source

1..988  
 /codon\_start=1  
 /product="cardiac junctin"  
 /protein\_id="AAK0614.1"  
 /db\_xref="GI:12655825"  
 /translation="MAEDKEAKGHKNGRRGGISGGSGFTFMVIALGLGWTISVAVV  
 WFDLVYEEVLKGLGVYDADGDGPDVDAKVLGGGLAKRTKAKAKEPIKEELK  
 KERGKAVPSKNEERQKKEQEDRGKRKPSDTSQKASAGKRDREKASDDKSS  
 KSKESWKAETKAVSKVAARDKDRGRSSGHAHVSKENGQKRKN"

ORIGIN

Query Match 1.3%; Score 31; DB 10; Length 919;  
 Best Local Similarity 100.0%; Pred. No. 0.00011;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 303 TTTGATGTGGATGATGCCAAAGTTTATTAG 333  
 Db 347 TTTGATGTGGATGATGCCAAAGTTTATTAG 377

RESULT 48

AF223414  
 LOCUS Mus musculus junctin-2 mRNA, complete cds. ROD 19-DEC-2002  
 DEFINITION Mus musculus junctin-2 mRNA, complete cds.  
 ACCESSION AF223414  
 VERSION AF223414.1 GI:27261147  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 988)  
 AUTHORS Hong, C.-S. and Kim, D.H.  
 TITLE Mouse junctin-2 mRNA  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 988)  
 AUTHORS Hong, C. and Kim, D.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-JAN-2000) Life Science, Kwangju Institute of Science and Technology (K-JIST), Puk-gu, Oryong-dong, Kwangju 500-712, South Korea

FEATURES

source

1..988  
 /codon\_start=1  
 /product="cardiac junctin"  
 /protein\_id="AAK0614.1"  
 /db\_xref="GI:12655825"  
 /translation="MAEDKEAKGHKNGRRGGISGGSGFTFMVIALGLGWTISVAVV  
 WFDLVYEEVLKGLGVYDADGDGPDVDAKVLGGGLAKRTKAKAKEPIKEELK  
 KERGKAVPSKNEERQKKEQEDRGKRKPSDTSQKASAGKRDREKASDDKSS  
 KSKESWKAETKAVSKVAARDKDRGRSSGHAHVSKENGQKRKN"

ORIGIN

Query Match 1.3%; Score 31; DB 10; Length 988;  
 Best Local Similarity 100.0%; Pred. No. 0.00011;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 303 TTTGATGTGGATGATGCCAAAGTTTATTAG 333  
 Db 382 TTTGATGTGGATGATGCCAAAGTTTATTAG 412

RESULT 49

AF223413  
 LOCUS Mus musculus junctin-1 mRNA, complete cds. ROD 19-DEC-2002  
 DEFINITION Mus musculus junctin-1 mRNA, complete cds.  
 ACCESSION AF223413

FEATURES

source

1..1013  
 /codon\_start=1  
 /product="cardiac junctin"  
 /protein\_id="AAK0614.1"  
 /db\_xref="GI:12655825"  
 /translation="MAEDKEAKGHKNGRRGGISGGSGFTFMVIALGLGWTISVAVV  
 WFDLVYEEVLKGLGVYDADGDGPDVDAKVLGGGLAKRTKAKAKEPIKEELK  
 KERGKAVPSKNEERQKKEQEDRGKRKPSDTSQKASAGKRDREKASDDKSS  
 KSKESWKAETKAVSKVAARDKDRGRSSGHAHVSKENGQKRKN"

AF223413.1 GI:27261145  
 Mus musculus (house mouse)  
 Mus musculus

REFERENCE 1 (bases 1 to 1013)  
 AUTHORS Hong, C. and Kim, D.H.  
 TITLE Mouse junctin-1 mRNA  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1013)  
 AUTHORS Hong, C. and Kim, D.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-JAN-2000) Life Science, Kwangju Institute of Science and Technology (K-JIST), Puk-gu, Oryong-dong, Kwangju 500-712, South Korea

FEATURES

source

1..1013  
 /codon\_start=1  
 /product="cardiac junctin"  
 /protein\_id="AAK0614.1"  
 /db\_xref="GI:12655825"  
 /translation="MAEDKEAKGHKNGRRGGISGGSGFTFMVIALGLGWTISVAVV  
 WFDLVYEEVLKGLGVYDADGDGPDVDAKVLGGGLAKRTKAKAKEPIKEELK  
 KERGKAVPSKNEERQKKEQEDRGKRKPSDTSQKASAGKRDREKASDDKSS  
 KSKESWKAETKAVSKVAARDKDRGRSSGHAHVSKENGQKRKN"

ORIGIN

Query Match 1.3%; Score 31; DB 10; Length 1013;  
 Best Local Similarity 100.0%; Pred. No. 0.00011;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 303 TTTGATGTGGATGATGCCAAAGTTTATTAG 333  
 Db 372 TTTGATGTGGATGATGCCAAAGTTTATTAG 402

RESULT 50

AF289490  
 LOCUS Mus musculus junctin mRNA, complete cds. ROD 18-DEC-2000  
 DEFINITION Mus musculus junctin mRNA, complete cds.  
 ACCESSION AF289490  
 VERSION AF289490.1 GI:11878117  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1733)  
 AUTHORS Dinchuk, J.E., Henderson, N.L., Burn, T.C., Huber, R., Ho, S.P., Link, J., O'Neil, K.T., Focht, R.J., Scully, M.S., Hollis, J.M., Hollis, G.F. and Friedman, P.A.  
 TITLE Aspartyl beta-hydroxylase (Asph) and an evolutionarily conserved isoform of Asph missing the catalytic domain share exons with junctin  
 JOURNAL J. Biol. Chem. 275 (50), 39543-39554 (2000)  
 MEDLINE 20564328  
 PUBMED 10956665  
 REFERENCE 2 (bases 1 to 1733)  
 AUTHORS Dinchuk, J.E., Henderson, N.L., Burn, T.C., Hollis, G.P. and Friedman, P.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (26-JUL-2000) Applied Biotechnology, DuPont Pharmaceuticals, P.O. Box 80400, Wilmington, DE 19880-0400, USA

FEATURES

source

1..1733  
 /codon\_start=1  
 /product="cardiac junctin"  
 /protein\_id="AAK0614.1"  
 /db\_xref="GI:12655825"  
 /translation="MAEDKEAKGHKNGRRGGISGGSGFTFMVIALGLGWTISVAVV  
 WFDLVYEEVLKGLGVYDADGDGPDVDAKVLGGGLAKRTKAKAKEPIKEELK  
 KERGKAVPSKNEERQKKEQEDRGKRKPSDTSQKASAGKRDREKASDDKSS  
 KSKESWKAETKAVSKVAARDKDRGRSSGHAHVSKENGQKRKN"

ORIGIN

Query Match 1.3%; Score 31; DB 10; Length 1013;  
 Best Local Similarity 100.0%; Pred. No. 0.00011;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 303 TTTGATGTGGATGATGCCAAAGTTTATTAG 333  
 Db 372 TTTGATGTGGATGATGCCAAAGTTTATTAG 402

RESULT 50

AF289490  
 LOCUS Mus musculus junctin mRNA, complete cds. ROD 18-DEC-2000  
 DEFINITION Mus musculus junctin mRNA, complete cds.  
 ACCESSION AF289490  
 VERSION AF289490.1 GI:11878117  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1733)  
 AUTHORS Dinchuk, J.E., Henderson, N.L., Burn, T.C., Huber, R., Ho, S.P., Link, J., O'Neil, K.T., Focht, R.J., Scully, M.S., Hollis, J.M., Hollis, G.F. and Friedman, P.A.  
 TITLE Aspartyl beta-hydroxylase (Asph) and an evolutionarily conserved isoform of Asph missing the catalytic domain share exons with junctin  
 JOURNAL J. Biol. Chem. 275 (50), 39543-39554 (2000)  
 MEDLINE 20564328  
 PUBMED 10956665  
 REFERENCE 2 (bases 1 to 1733)  
 AUTHORS Dinchuk, J.E., Henderson, N.L., Burn, T.C., Hollis, G.P. and Friedman, P.A.  
 TITLE Direct Submission  
 JOURNAL Submitted (26-JUL-2000) Applied Biotechnology, DuPont Pharmaceuticals, P.O. Box 80400, Wilmington, DE 19880-0400, USA

FEATURES

source

1..1733  
 /codon\_start=1  
 /product="cardiac junctin"  
 /protein\_id="AAK0614.1"  
 /db\_xref="GI:12655825"  
 /translation="MAEDKEAKGHKNGRRGGISGGSGFTFMVIALGLGWTISVAVV  
 WFDLVYEEVLKGLGVYDADGDGPDVDAKVLGGGLAKRTKAKAKEPIKEELK  
 KERGKAVPSKNEERQKKEQEDRGKRKPSDTSQKASAGKRDREKASDDKSS  
 KSKESWKAETKAVSKVAARDKDRGRSSGHAHVSKENGQKRKN"

ORIGIN

Query Match 1.3%; Score 31; DB 10; Length 1013;  
 Best Local Similarity 100.0%; Pred. No. 0.00011;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 303 TTTGATGTGGATGATGCCAAAGTTTATTAG 333  
 Db 382 TTTGATGTGGATGATGCCAAAGTTTATTAG 412

RESULT 49

AF223413  
 LOCUS Mus musculus junctin-1 mRNA, complete cds. ROD 19-DEC-2002  
 DEFINITION Mus musculus junctin-1 mRNA, complete cds.  
 ACCESSION AF223413

```

Db      251 TTTGATGTGGATGATGCCAAAGTTTATTAG 281

RESULT 52
AF302654      2637 bp      mRNA      linear      ROD 29-OCT-2001
LOCUS      Mus musculus cardiac junctate 2 mRNA, complete cds.
DEFINITION
ACCESSION      AF302654
VERSION      AF302654.1 GI:15824391
KEYWORDS
SOURCE      Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 2637)
AUTHORS      Hong, C.-S. and Kim, D.H.
TITLE      Cloning of mouse junctin homologs
JOURNAL      Unpublished
REFERENCE
2 (bases 1 to 2637)
AUTHORS      Hong, C.-S. and Kim, D.H.
TITLE      Direct Submission
JOURNAL      Submitted (05-SEP-2000) Life Science, Kwangju Institute of Science
and Technology, Puk-gu, Oryong-dong, Kwangju 500-712, South Korea
FEATURES
source
1..2637
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
120..899
/codon_start=1
/product="cardiac junctate 2"
/protein_id="AAL09320.1"
/db_xref="GI:15824392"
/translation="MAEDKEAKHGKNGRRGSGSFTFMVIALLGWTSVAVV
WFDLVYEEVLKGLGYDADGDGFDVDDAKVLLEGGGLAKRTAKAKEPKEIK
TFPEAEHAELEQAEAGLEADGAGEPQVEEDFLTVTSDRDFLEPTVHE
EIEDTVHEDTASQNHNDMEETNEQNSDPSEAVTDAGVLLPHAEVRHQDYDEPV
YEPSEHGEVSEISNTIDSSIISEEINVASVEEQDTPPT"

CDS
Query Match      1.3%; Score 31; DB 10; Length 2637;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      303 TTTGATGTGGATGATGCCAAAGTTTATTAG 333
|||||
Db      324 TTTGATGTGGATGATGCCAAAGTTTATTAG 354

RESULT 53
AC143465/c      158289 bp      DNA      linear      HTG 09-APR-2003
LOCUS      Macaca mulatta clone CH250-269N19, *** SEQUENCING IN PROGRESS ***.
DEFINITION
ACCESSION      AC143465
VERSION      AC143465.1 GI:29648741
KEYWORDS      HTG; HTGS PHASE2; HTGS_PGI.
SOURCE      Macaca mulatta (rhesus monkey)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE
1 (bases 1 to 158289)
AUTHORS      Csuros, M. and Milosavljevic, A.
TITLE      Pooled genomic indexing (PGI): mathematical analysis and experiment
design
JOURNAL      (in) Guigo, R. and Gusfield, D. (Eds.),
ALGORITHMS IN BIOINFORMATICS, SECOND INTERNATIONAL WORKSHOP, WABI
2002, ROME, ITALY, SEPTEMBER 17-21, 2002, PROCEEDINGS: 10-28;
Springer (2002)
REFERENCE
2 (bases 1 to 158289)
AUTHORS      Milosavljevic, A., Sodergren, E., Csuros, M., Li, B., Jackson, A.R.,
Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C., Alsbrooks, S.L.,

Db      251 TTTGATGTGGATGATGCCAAAGTTTATTAG 281

RESULT 52
AF302654      2637 bp      mRNA      linear      ROD 29-OCT-2001
LOCUS      Mus musculus cardiac junctate 2 mRNA, complete cds.
DEFINITION
ACCESSION      AF302654
VERSION      AF302654.1 GI:15824391
KEYWORDS
SOURCE      Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 2637)
AUTHORS      Hong, C.-S. and Kim, D.H.
TITLE      Cloning of mouse junctin homologs
JOURNAL      Unpublished
REFERENCE
2 (bases 1 to 2637)
AUTHORS      Hong, C.-S. and Kim, D.H.
TITLE      Direct Submission
JOURNAL      Submitted (05-SEP-2000) Life Science, Kwangju Institute of Science
and Technology, Puk-gu, Oryong-dong, Kwangju 500-712, South Korea
FEATURES
source
1..2637
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
120..899
/codon_start=1
/product="cardiac junctate 2"
/protein_id="AAL09320.1"
/db_xref="GI:15824392"
/translation="MAEDKEAKHGKNGRRGSGSFTFMVIALLGWTSVAVV
WFDLVYEEVLKGLGYDADGDGFDVDDAKVLLEGGGLAKRTAKAKEPKEIK
TFPEAEHAELEQAEAGLEADGAGEPQVEEDFLTVTSDRDFLEPTVHE
EIEDTVHEDTASQNHNDMEETNEQNSDPSEAVTDAGVLLPHAEVRHQDYDEPV
YEPSEHGEVSEISNTIDSSIISEEINVASVEEQDTPPT"

CDS
Query Match      1.3%; Score 31; DB 10; Length 1733;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      303 TTTGATGTGGATGATGCCAAAGTTTATTAG 333
|||||
Db      329 TTTGATGTGGATGATGCCAAAGTTTATTAG 359

RESULT 51
AF302655      2432 bp      mRNA      linear      ROD 29-OCT-2001
LOCUS      Mus musculus cardiac junctate 3 mRNA, complete cds.
DEFINITION
ACCESSION      AF302655
VERSION      AF302655.1 GI:15824393
KEYWORDS
SOURCE      Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 2432)
AUTHORS      Hong, C.-S. and Kim, D.H.
TITLE      Cloning of mouse junctin homologs
JOURNAL      Unpublished
REFERENCE
2 (bases 1 to 2432)
AUTHORS      Hong, C.-S. and Kim, D.H.
TITLE      Direct Submission
JOURNAL      Submitted (05-SEP-2000) Life Science, Kwangju Institute of Science
and Technology, Puk-gu, Oryong-dong, Kwangju 500-712, South Korea
FEATURES
source
1..2432
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
47..694
/codon_start=1
/product="cardiac junctate 3"
/protein_id="AAL09321.1"
/db_xref="GI:15824394"
/translation="MAEDKEAKHGKNGRRGSGSFTFMVIALLGWTSVAVV
WFDLVYEEVLKGLGYDADGDGFDVDDAKVLLEGGGLAKRTAKAKEPKEIK
DDRFLEPTVHEEDTVHEDTASQNHNDMEETNEQNSDPSEAVTDAGVLLP
HAEVRHQDYDEPVPEPSHEGVEISNTIDSSIISEEINVASVEEQDTPPT"

CDS
Query Match      1.3%; Score 31; DB 10; Length 2432;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      303 TTTGATGTGGATGATGCCAAAGTTTATTAG 333
|||||

```

Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbara, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.P., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davis, C., Davy-Carroll, E., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douchwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Egan, A., Earnhart, C., Edwards, C.C., Elhaj, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Foster, P., Frantz, P., Gabis, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorell, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Ioshikhes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Ma, J., Maheshwari, M., Mapua, P., Marondel, I., Martin, R., Martindale, A., Martinez, E., Masey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Merscher, S., Metzker, M., Miller, A., Miner, G., Morris, Z., Mitchell, T., Mohabbat, K., Montgomery, K.T., Morgan, M., Nguyen, A., Nguyen, N., Oragunye, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scher, S., Scott, G., Shen, H., Shim, C., Shooshkari, N., Sison, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabot, P., Tamerisa, A., Tameria, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Uman, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczky, R., Wooten, S., Worley, K., Wu, C., Wu, Y.F., Zhou, J., Zorilla, S., Zucherlapati, R., Weinstock, G. and Gibbs, R.

Direct Submission  
Unpublished  
3 (bases 1 to 158289)  
Worley, K.C.

Submitted (09-APR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

----- Genome Center -----  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

----- Project Information -----  
Center project name: LBGJ  
Center clone name: CH250-269N19

----- Summary Statistics -----  
Chemistry: Dye-terminator  
Dye-terminator: Big Dye, inf% of reads  
Consensus quality: 7790 bases at least Q40  
Consensus quality: 9211 bases at least Q30  
Consensus quality: 10542 bases at least Q20

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html))  
\* NOTE: the contigs are based on the application  
of the FGI method using the Human genome (NCBI build 31)  
as the comparative genome.  
\* NOTE: This is a 'working draft' sequence. It currently  
consists of 1 contigs. Gaps between the contigs  
are represented as runs of N. The order of the pieces  
is believed to be correct as given, however the sizes  
of the gaps between them are based on estimates that have

\* Provided by the submittor.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 158289: contig of 158289 bp in length.

## FEATURES

Location/Qualifiers  
1..158289  
/organism="Macaca mulatta"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9544"  
/clones="CH250-269N19" -  
1..158289  
/note="assembly name:CH250-269N19.1B  
CONFIDENCES:\_0.83"

## misc\_feature

## ORIGIN

Query Match 1.3%; Score 30; DB 2; Length 158289;  
Best Local Similarity 100.0%; Pred. No. 0.00043;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Cy 1310 AGGTCATATGAGAGGTTCCCTGCTTACCCT 1339  
Db 132122 AGGTCATATGAGAGGTTCCCTGCTTACCCT 132093

## RESULT 54

AC095635/c  
LOCUS AC095635 234548 bp DNA linear HTG 09-MAY-2003  
Rattus norvegicus clone CH230-8K1, WORKING DRAFT SEQUENCE, 2  
unordered pieces.

## ACCESSION

AC095635 GI:30467118

HTG: HTGS PHASE1; HTGS DRAFT; HTGS\_FULLTOP.

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

1 (bases 1 to 234548)  
Muzny, D., Maric, E., Metzker, M., Lee, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandavani, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, P., Fraser, C.M., Gabis, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvay, P., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Loulseghe, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Munidasa, M., Murphy, M., Nair, L., Morgan, M., Morris, K., Morris, N., Nguyen, N., Norris, S., Nankervis, C., Neal, D., Newton, N., Okwuonu, G., Olarpunsgoon, A., Fal, S., Parks, K., Nwokenkwo, O., Okwuonu, G., Olarpunsgoon, A., Fal, S., Parks, K., Pasternak, S., Paul, H., Perez, L., Prannkoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.L.,



```

Paazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rockey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villanasa,D., Waldron,L., Walker,B., Wang,D.,
Wang,Q., Wang,S., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wlezyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,J., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhauser,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 234548)
Worley,K.C.
Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 234548)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 9, 2003 this sequence version replaced gi:24817791.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GCWM
Center clone name: CH230-8K1
----- Summary Statistics
Assembly program: Atlas
Consensus quality: 224348 bases at least Q40
Consensus quality: 224642 bases at least Q30
Consensus quality: 225972 bases at least Q20
Estimated insert size: 232080; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
consists of 2 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
* 1 232993: contig of 232993 bp in length
* 232994 233093: gap of unknown length
* 233094 234548: contig of 1455 bp in length.
Location/Qualifiers
1. .234548

```

```

/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-8K1"
1. .1711
/misc_feature
/notes="wgs_contig"

ORIGIN

Query Match 1.2%; Score 29; DB 2; Length 234548;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1839 GGTCTTCTTCCTGCTGAGATGAAACCT 1867
|||||
DB 175547 GGTCTTCTTCCTGCTGAGATGAAACCT 175519

RESULT 55
AC128056/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-16022, WORKING DRAFT SEQUENCE, 4
unordered pieces.
AC128056
AC128056.3 GI-25007693
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 275547)
Muzny,D.,Marie., Metzker,M., Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Ayagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Cesar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., De Souza,L.,
Davila,M., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Khan,Z., King,L., Kovat,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwackeleme,O., Okwou,G., Olarpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,

```

Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

# TITLE JOURNAL

REFERENCE  
AUTHORS

TITLE  
JOURNAL

Submitted (19-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 275547)  
Rat Genome Sequencing Consortium.

# REFERENCE AUTHORS

TITLE  
JOURNAL

Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Nov 15, 2002 this sequence version replaced gi:23665157.

# COMMENT

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Center: Genome Center

Center: Baylor College of Medicine

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

----- Project Information

Center project name: KAPS

Center clone name: CH230-16022

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 223348 bases at least Q40

Consensus quality: 225892 bases at least Q30

Consensus quality: 227931 bases at least Q20

Estimated insert size: 224883; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

-----

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 4 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence.  
\* As soon as it is available and the accession number will  
\* be preserved.

\* 1 269305: contig of 269305 bp in length

\* 269306 269405: gap of unknown length

\* 269406 273098: contig of 3693 bp in length

\* 273099 273198: gap of unknown length

\* 273199 274358: contig of 1160 bp in length

\* 274359 274459: gap of unknown length

\* 274459 275547: contig of 1089 bp in length.

Location/Qualifiers

1..275547

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-16022"

# FEATURES

source

misc\_feature  
1..1780  
/note="wgs contig"  
misc\_feature  
266655..267994  
/note="wgs contig"  
misc\_feature  
268045..269305  
/note="wgs contig"

# ORIGIN

Query Match 1.2%; Score 29; DB 2; Length 275547;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1839 GGTCTCTTCTCTGCTGCTGAGGATGAAACCT 1867

|||||

Db 28866 GGTCTCTTCTCTGCTGCTGAGGATGAAACCT 28838

RESULT 56

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..27

/organism="synthetic construct"

/mol\_type="unassigned DNA"

/db\_xref="taxon:32630"

/note="QPCR primer"

# ORIGIN

Query Match 1.2%; Score 27; DB 6; Length 27;  
Best Local Similarity 100.0%; Pred. No. 0.018;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 286 CTGATGCTGATGAGAGATTTTGATGTGG 312

|||||

Db 27 CTGATGCTGATGAGAGATTTTGATGTGG 1

RESULT 57

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SEGMENT

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

1..275547

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/clone="CH230-16022"

# FEATURES

source

H010058505  
Mus musculus aspartyl beta-hydroxylase (ASPH) gene, exon 15.  
AF289209  
AF289209.1 GI:11692638  
5 of 11  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 2013)  
Dinchuk, J.E., Henderson, N.L., Burn, T.C., Huber, R., Ho, S.P.,  
Link, J., O'Neil, K.T., Focht, R.J., Scully, M.S., Hollis, J.M.,  
Hollis, G.F. and Friedman, P.A.  
Aspartyl beta -hydroxylase (Asph) and an evolutionarily conserved  
isoform of Asph missing the catalytic domain share exons with  
junction  
J. Biol. Chem. 275 (50), 39543-39554 (2000)  
20564328  
10956665  
2 (bases 1 to 2013)  
Henderson, N.L., Dinchuk, J.E., Burn, T.C., Hollis, G.F. and

Friedman, P.A.  
 Direct Submission  
 Submitted (25-JUL-2000) Applied Biotechnology, DuPont  
 Pharmaceuticals Company, PO BOX 80400, Wilmington, DE 19880, USA  
 Location/Qualifiers  
 1. 2013  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 292..378  
 /gene="ASPH"  
 /number=15

exon

ORIGIN

Query Match 1.2%; Score 27; DB 10; Length 1013;  
 Best Local Similarity 100.0%; Pred. No. 0.019;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1077 AAAATTGAGGAGCAGTCAATGCATT 1103  
 |||||  
 DB 295 AAAATTGAGGAGCAGTCAATGCATT 321

RESULT 58  
 AC096178 252366 bp DNA linear HTG 10-MAY-2003  
 LOCUS Rattus norvegicus clone CH230-11D6, WORKING DRAFT SEQUENCE, 3  
 DEFINITION unordered pieces.  
 AC096178  
 AC096178.6 GI:30522511  
 HTG: HTGS PHASE1; HTGS DRAFT; HTGS\_FULLTOP.  
 VERSION  
 KEYWORDS Rattus norvegicus (Norway rat)  
 SOURCE  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (bases 1 to 252366)  
 Muzny, D., Marie, M., Metzker, M., Lee, A., Abramson, S., Adams, C., Alder, J.,  
 Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,  
 Anyalebechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H.,  
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,  
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
 Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinu, H., Divye, K.,  
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K.,  
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,  
 Fernandez, S., Finley, M., Flag, N., Forbes, L., Foster, M., Foster, P.,  
 Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
 Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,  
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,  
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M.,  
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,  
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
 Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
 Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
 Lorenshehwa, L., Loulseghe, H., Lozado, R.J., Lu, X., Ma, J.,  
 Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,  
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,  
 Mathew, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,  
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
 Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,  
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
 Nwokenne, O., Okwuon, G., Olarnpungsoo, A., Pal, S., Parks, K.,  
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,  
 Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, L.-L.,  
 Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,  
 Reilly, B., Reilly, W., Ren, Y., Reuter, M., Richards, S., Riggs, P.,

Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Shen, H.,  
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,  
 Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,  
 Sneed, A., Sodergerm, E., Song, X.-Z., Sorelle, R., Sosa, J.,  
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,  
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,  
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,  
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,  
 Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K.,  
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,  
 Weinstock, G., and Gibbs, R.A.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 252366)  
 Worley, K.C.  
 Direct Submission  
 Submitted (17-SEP-2001) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 252366)  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (10-MAY-2003) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On May 10, 2003 this sequence version replaced gi:24818276.  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Project name: GEPB  
 Center project name: CH230-11D6  
 Center clone name: CH230-11D6  
 ----- Summary Statistics  
 Assembly program: Atlas 3.0;  
 Consensus quality: 230832 bases at least Q40  
 Consensus quality: 233984 bases at least Q30  
 Consensus quality: 235460 bases at least Q20  
 Estimated insert size: 242518; sum-of-contigs estimation  
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 3 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 248556: contig of 248556 bp in length  
 \* 248557 248556: gap of unknown length  
 \* 24857 250653: contig of 2007 bp in length  
 \* 250664 250763: gap of unknown length  
 \* 250764 252366: contig of 1603 bp in length.  
 \* Location/Qualifiers  
 1. 252366

FEATURES  
 source

```

/misc_feature
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-11D6"
264. .906
/notes="clone_boundary
site:EcoRI
end_sequence:BH340110"
misc_feature
20540. .40522
/notes="clone_boundary
clone end:Sp6
site:EcoRI
end_sequence:BH340112"
ORIGIN
Query Match 1.1%; Score 27; DB 2; Length 252366;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 452 AAAGAACAAATTCAGTCCCTTCTCCA 478
DB 170125 AAAGAACAAATTCAGTCCCTTCTCCA 170151
RESULT 59
H01005807 6629 bp DNA linear ROD 13-MAR-2001
LOCUS
DEFINITION
Mus musculus aspartyl beta-hydroxylase (ASPH) gene, exons 18 and
19.
ACCESSION AF289211
VERSION AF289211.1 GI:11692640
KEYWORDS
SEGMENT
SOURCE
7 of 11
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 6629)
Dinchuk,J.E., Henderson,N.L., Burn,T.C., Huber,R., Ho,S.P.,
Link,J., O'Neil,K.T., Focht,R.J., Scully,M.S., Hollis,J.M.,
Hollis,G.F. and Friedman,P.A.
TITLE
Aspartyl beta -hydroxylase (Asph) and an evolutionarily conserved
isoform of Asph missing the catalytic domain share exons with
junctin
JOURNAL
MEDLINE 20564328
PUBMED 10956665
REFERENCE
2 (bases 1 to 6629)
Henderson,N.L., Dinchuk,J.E., Burn,T.C., Hollis,G.F. and
Friedman,P.A.
TITLE
Direct Submission
JOURNAL
Submitted (25-JUL-2000) Applied Biotechnology, DuPont
Pharmaceuticals Company, PO BOX 80400, Wilmington, DE 19880, USA
FEATURES
Source
1. .6629
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
1764. .1862
/genes="ASPH"
/number=18
exon
4834. .4923
/genes="ASPH"
/number=19
exon
ORIGIN
Query Match 1.1%; Score 26; DB 10; Length 6629;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1494 GGCTTCATCTGAGGCACAGACAA 1519
DB 1809 GGCTTCATCTGAGGCACAGACAA 1834
RESULT 60
H01005810 12095 bp DNA linear ROD 13-MAR-2001
LOCUS
DEFINITION
Mus musculus aspartyl beta-hydroxylase (ASPH) gene, exons 22 and
23.
ACCESSION AF289214
VERSION AF289214.1 GI:11692643
KEYWORDS
SEGMENT
SOURCE
10 of 11
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 12095)
Dinchuk,J.E., Henderson,N.L., Burn,T.C., Huber,R., Ho,S.P.,
Link,J., O'Neil,K.T., Focht,R.J., Scully,M.S., Hollis,J.M.,
Hollis,G.F. and Friedman,P.A.
TITLE
Aspartyl beta -hydroxylase (Asph) and an evolutionarily conserved
isoform of Asph missing the catalytic domain share exons with
junctin
JOURNAL
MEDLINE 20564328
PUBMED 10956665
REFERENCE
2 (bases 1 to 12095)
Henderson,N.L., Dinchuk,J.E., Burn,T.C., Hollis,G.F. and
Friedman,P.A.
TITLE
Direct Submission
JOURNAL
Submitted (25-JUL-2000) Applied Biotechnology, DuPont
Pharmaceuticals Company, PO BOX 80400, Wilmington, DE 19880, USA
FEATURES
Source
1. .12095
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
5360. .5451
/genes="ASPH"
/number=22
exon
5805. .5938
/genes="ASPH"
/number=23
exon
ORIGIN
Query Match 1.1%; Score 26; DB 10; Length 12095;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2055 CCCACAACTGCAGGCTCGAATGCA 2080
DB 5856 CCCACAACTGCAGGCTCGAATGCA 5881
RESULT 61
AF016503/c
LOCUS
DEFINITION
Rattus norvegicus procollagen C-proteinase enhancer protein (PCPE)
mRNA, complete cds.
ACCESSION AF016503
VERSION AF016503.1 GI:4102818
KEYWORDS
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 1519)
Masuda,M., Igarashi,H., Kano,M. and Yoshikura,H.
Provitral integration into the Procollagen C-proteinase Enhancer
Protein Gene and Its Effects in Cultured Rat Fibroblasts Revealed
by an Excisable 'Hit-and-Run' Retroviral Vector
Unpublished
JOURNAL

```

2 (bases 1 to 1519)  
Masuda, M. and Igarashi, H.  
Direct Submission  
Submitted (30-JUL-1997) Department of Microbiology, Graduate School  
of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo  
113, Japan

Location/Qualifiers  
1. .1519  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="F344"  
/db\_xref="taxon:10116"  
/cell\_line="Rat2"  
/cell\_type="fibroblast"  
/dev\_stage="embryo"  
/gene="PCPE"  
1. .1519  
56. .1462  
/codon="PCPE"  
/codon\_start=1  
/product="procollagen C-proteinase enhancer protein"  
/protein\_id="AAD01592.1"  
/db\_xref="GI:4102819"  
/translation="MLPAALTSLLGLPFLAWLPLARGQTPNTRPVFLCGDVTCES  
GVVAGEPNLYPPNKKICITVPEBQTVLSRFVDMELHPSCRYDALEVPAGSGT  
SQRLQAGCTGPRPAPVAPGNQVLTMTDEGTGGRGFLWYSGRATSGTEHFCGG  
NKEAAGCTLTTPNPESDYPGICSCWHIIAPSNQVIMLTFGKFDVEPDPTCRYDSVS  
VNGAVSDSKRLGKFCGDKAPSPISSEGNELLQVFSDLSTADGFSASRYRLPDA  
VKEASPGEDAQHGQPSRDPKTGTGPKVPKQVPEKPGSPATQATPVAP  
DAPSTCPQKYKRGSTLQNSCSSLWVGTGKAMVWGREGLTWTVSLGVYKTGDL  
DLPSPATSLKPYPCQMPMKKGASYLLMGQVEENRGFPILPPSFVVLVYRPNDQD  
ILNLSKRCPCSPRPDA"

Query Match 1.0%; Score 24; DB 10; Length 1519;  
Best Local Similarity 100.0%; Pred. No. 0.83; 0; Indels 0; Gaps 0;  
Matches 24; Conservative 0; Mismatches 0;

QY 36 AGCAGCGGCAACAGCAGCAGCAGC 59  
|||||  
DB 29 AGCAGCGGCAACAGCAGCAGCAGC 6  
|||||

ORIGIN  
Query Match 1.0%; Score 24; DB 10; Length 1519;  
Best Local Similarity 100.0%; Pred. No. 0.83; 0; Indels 0; Gaps 0;  
Matches 24; Conservative 0; Mismatches 0;

QY 36 AGCAGCGGCAACAGCAGCAGCAGC 59  
|||||  
DB 29 AGCAGCGGCAACAGCAGCAGCAGC 6  
|||||

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1. .1519  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="F344"  
/db\_xref="taxon:10116"  
/cell\_line="Rat2"  
/cell\_type="fibroblast"  
/dev\_stage="embryo"  
/gene="PCPE"  
1. .1519  
56. .1462  
/codon="PCPE"  
/codon\_start=1  
/product="procollagen C-proteinase enhancer protein"  
/protein\_id="AAD01592.1"  
/db\_xref="GI:4102819"  
/translation="MLPAALTSLLGLPFLAWLPLARGQTPNTRPVFLCGDVTCES  
GVVAGEPNLYPPNKKICITVPEBQTVLSRFVDMELHPSCRYDALEVPAGSGT  
SQRLQAGCTGPRPAPVAPGNQVLTMTDEGTGGRGFLWYSGRATSGTEHFCGG  
NKEAAGCTLTTPNPESDYPGICSCWHIIAPSNQVIMLTFGKFDVEPDPTCRYDSVS  
VNGAVSDSKRLGKFCGDKAPSPISSEGNELLQVFSDLSTADGFSASRYRLPDA  
VKEASPGEDAQHGQPSRDPKTGTGPKVPKQVPEKPGSPATQATPVAP  
DAPSTCPQKYKRGSTLQNSCSSLWVGTGKAMVWGREGLTWTVSLGVYKTGDL  
DLPSPATSLKPYPCQMPMKKGASYLLMGQVEENRGFPILPPSFVVLVYRPNDQD  
ILNLSKRCPCSPRPDA"

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1. .1519  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="F344"  
/db\_xref="taxon:10116"  
/cell\_line="Rat2"  
/cell\_type="fibroblast"  
/dev\_stage="embryo"  
/gene="PCPE"  
1. .1519  
56. .1462  
/codon="PCPE"  
/codon\_start=1  
/product="procollagen C-proteinase enhancer protein"  
/protein\_id="AAD01592.1"  
/db\_xref="GI:4102819"  
/translation="MLPAALTSLLGLPFLAWLPLARGQTPNTRPVFLCGDVTCES  
GVVAGEPNLYPPNKKICITVPEBQTVLSRFVDMELHPSCRYDALEVPAGSGT  
SQRLQAGCTGPRPAPVAPGNQVLTMTDEGTGGRGFLWYSGRATSGTEHFCGG  
NKEAAGCTLTTPNPESDYPGICSCWHIIAPSNQVIMLTFGKFDVEPDPTCRYDSVS  
VNGAVSDSKRLGKFCGDKAPSPISSEGNELLQVFSDLSTADGFSASRYRLPDA  
VKEASPGEDAQHGQPSRDPKTGTGPKVPKQVPEKPGSPATQATPVAP  
DAPSTCPQKYKRGSTLQNSCSSLWVGTGKAMVWGREGLTWTVSLGVYKTGDL  
DLPSPATSLKPYPCQMPMKKGASYLLMGQVEENRGFPILPPSFVVLVYRPNDQD  
ILNLSKRCPCSPRPDA"

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1. .1519  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="F344"  
/db\_xref="taxon:10116"  
/cell\_line="Rat2"  
/cell\_type="fibroblast"  
/dev\_stage="embryo"  
/gene="PCPE"  
1. .1519  
56. .1462  
/codon="PCPE"  
/codon\_start=1  
/product="procollagen C-proteinase enhancer protein"  
/protein\_id="AAD01592.1"  
/db\_xref="GI:4102819"  
/translation="MLPAALTSLLGLPFLAWLPLARGQTPNTRPVFLCGDVTCES  
GVVAGEPNLYPPNKKICITVPEBQTVLSRFVDMELHPSCRYDALEVPAGSGT  
SQRLQAGCTGPRPAPVAPGNQVLTMTDEGTGGRGFLWYSGRATSGTEHFCGG  
NKEAAGCTLTTPNPESDYPGICSCWHIIAPSNQVIMLTFGKFDVEPDPTCRYDSVS  
VNGAVSDSKRLGKFCGDKAPSPISSEGNELLQVFSDLSTADGFSASRYRLPDA  
VKEASPGEDAQHGQPSRDPKTGTGPKVPKQVPEKPGSPATQATPVAP  
DAPSTCPQKYKRGSTLQNSCSSLWVGTGKAMVWGREGLTWTVSLGVYKTGDL  
DLPSPATSLKPYPCQMPMKKGASYLLMGQVEENRGFPILPPSFVVLVYRPNDQD  
ILNLSKRCPCSPRPDA"

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1. .1519  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="F344"  
/db\_xref="taxon:10116"  
/cell\_line="Rat2"  
/cell\_type="fibroblast"  
/dev\_stage="embryo"  
/gene="PCPE"  
1. .1519  
56. .1462  
/codon="PCPE"  
/codon\_start=1  
/product="procollagen C-proteinase enhancer protein"  
/protein\_id="AAD01592.1"  
/db\_xref="GI:4102819"  
/translation="MLPAALTSLLGLPFLAWLPLARGQTPNTRPVFLCGDVTCES  
GVVAGEPNLYPPNKKICITVPEBQTVLSRFVDMELHPSCRYDALEVPAGSGT  
SQRLQAGCTGPRPAPVAPGNQVLTMTDEGTGGRGFLWYSGRATSGTEHFCGG  
NKEAAGCTLTTPNPESDYPGICSCWHIIAPSNQVIMLTFGKFDVEPDPTCRYDSVS  
VNGAVSDSKRLGKFCGDKAPSPISSEGNELLQVFSDLSTADGFSASRYRLPDA  
VKEASPGEDAQHGQPSRDPKTGTGPKVPKQVPEKPGSPATQATPVAP  
DAPSTCPQKYKRGSTLQNSCSSLWVGTGKAMVWGREGLTWTVSLGVYKTGDL  
DLPSPATSLKPYPCQMPMKKGASYLLMGQVEENRGFPILPPSFVVLVYRPNDQD  
ILNLSKRCPCSPRPDA"

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1. .1519  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="F344"  
/db\_xref="taxon:10116"  
/cell\_line="Rat2"  
/cell\_type="fibroblast"  
/dev\_stage="embryo"  
/gene="PCPE"  
1. .1519  
56. .1462  
/codon="PCPE"  
/codon\_start=1  
/product="procollagen C-proteinase enhancer protein"  
/protein\_id="AAD01592.1"  
/db\_xref="GI:4102819"  
/translation="MLPAALTSLLGLPFLAWLPLARGQTPNTRPVFLCGDVTCES  
GVVAGEPNLYPPNKKICITVPEBQTVLSRFVDMELHPSCRYDALEVPAGSGT  
SQRLQAGCTGPRPAPVAPGNQVLTMTDEGTGGRGFLWYSGRATSGTEHFCGG  
NKEAAGCTLTTPNPESDYPGICSCWHIIAPSNQVIMLTFGKFDVEPDPTCRYDSVS  
VNGAVSDSKRLGKFCGDKAPSPISSEGNELLQVFSDLSTADGFSASRYRLPDA  
VKEASPGEDAQHGQPSRDPKTGTGPKVPKQVPEKPGSPATQATPVAP  
DAPSTCPQKYKRGSTLQNSCSSLWVGTGKAMVWGREGLTWTVSLGVYKTGDL  
DLPSPATSLKPYPCQMPMKKGASYLLMGQVEENRGFPILPPSFVVLVYRPNDQD  
ILNLSKRCPCSPRPDA"

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1. .1519  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="F344"  
/db\_xref="taxon:10116"  
/cell\_line="Rat2"  
/cell\_type="fibroblast"  
/dev\_stage="embryo"  
/gene="PCPE"  
1. .1519  
56. .1462  
/codon="PCPE"  
/codon\_start=1  
/product="procollagen C-proteinase enhancer protein"  
/protein\_id="AAD01592.1"  
/db\_xref="GI:4102819"  
/translation="MLPAALTSLLGLPFLAWLPLARGQTPNTRPVFLCGDVTCES  
GVVAGEPNLYPPNKKICITVPEBQTVLSRFVDMELHPSCRYDALEVPAGSGT  
SQRLQAGCTGPRPAPVAPGNQVLTMTDEGTGGRGFLWYSGRATSGTEHFCGG  
NKEAAGCTLTTPNPESDYPGICSCWHIIAPSNQVIMLTFGKFDVEPDPTCRYDSVS  
VNGAVSDSKRLGKFCGDKAPSPISSEGNELLQVFSDLSTADGFSASRYRLPDA  
VKEASPGEDAQHGQPSRDPKTGTGPKVPKQVPEKPGSPATQATPVAP  
DAPSTCPQKYKRGSTLQNSCSSLWVGTGKAMVWGREGLTWTVSLGVYKTGDL  
DLPSPATSLKPYPCQMPMKKGASYLLMGQVEENRGFPILPPSFVVLVYRPNDQD  
ILNLSKRCPCSPRPDA"

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1. .1519  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="F344"  
/db\_xref="taxon:10116"  
/cell\_line="Rat2"  
/cell\_type="fibroblast"  
/dev\_stage="embryo"  
/gene="PCPE"  
1. .1519  
56. .1462  
/codon="PCPE"  
/codon\_start=1  
/product="procollagen C-proteinase enhancer protein"  
/protein\_id="AAD01592.1"  
/db\_xref="GI:4102819"  
/translation="MLPAALTSLLGLPFLAWLPLARGQTPNTRPVFLCGDVTCES  
GVVAGEPNLYPPNKKICITVPEBQTVLSRFVDMELHPSCRYDALEVPAGSGT  
SQRLQAGCTGPRPAPVAPGNQVLTMTDEGTGGRGFLWYSGRATSGTEHFCGG  
NKEAAGCTLTTPNPESDYPGICSCWHIIAPSNQVIMLTFGKFDVEPDPTCRYDSVS  
VNGAVSDSKRLGKFCGDKAPSPISSEGNELLQVFSDLSTADGFSASRYRLPDA  
VKEASPGEDAQHGQPSRDPKTGTGPKVPKQVPEKPGSPATQATPVAP  
DAPSTCPQKYKRGSTLQNSCSSLWVGTGKAMVWGREGLTWTVSLGVYKTGDL  
DLPSPATSLKPYPCQMPMKKGASYLLMGQVEEN

```

/product="Txk"
/protein_id="AAC52264.1"
/db_xref="GI:562125"
/translation="MILSSYSSFQSVLCCCRCSVQKROVTRQISSLREBELSEKHS
ORQRPWFAKMGKTQSNRGVQPSRKKEPLPQPPDERIQVAKLYDFLPREPGLA
LKRAEVLILERCOPHWKARDFGNEGLIFSNVTENRNLANLEIYWHKNTIRNOT
ERLQKAKAGATIVDRSHLGSYITISVTRARRHTQSSIKHYQIKNDGQWYITER
HLFSPVELIOYHOYNAAGLISRLRYPIGLLGSCLPATSGFSYKEMIDPSLAFVKE
IGSQFGVHLGSEWRAHIPVAIKAINEGMSEEDFIEAKVMKLSHSLVLYGVC
OQKPLYITVTFMGCLLDYLRERGOQLKALLSMCODICEGMAYLERSCVIHRDLA
ARNCVSSACVKSIDFGMARYVLDDEVYISSGAKFPVKMCPPEVPHFNKYSKSDVM
SFGLVMEVTEGKMPFENKSNLVQVVEAISQGFRLYRPHLAPMTIYRVVMSCHWSPK
GRPTFAELLQVLTEIAETW"
1633..2204

3'UTR
ORIGIN
Query Match 1.0%; Score 24; DB 10; Length 2204;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 GAGCAGCGGCAACGACGACGACGAG 58
|||||
Db 110 GAGCAGCGGCAACGACGACGACGAG 87

RESULT 64
MUSRLK/c
LOCUS Mus cckii resting lymphocyte kinase (RLK) gene, complete cds.
DEFINITION
ACCESSION L35268
VERSION L35268.1 GI:623442
KEYWORDS protein tyrosine kinase; resting lymphocyte kinase.
SOURCE Mus cckii (Cook's mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Hu.O., Davidson,D., Schwartzberg,P.L., Macchiarini,F.,
Lenard,M.J., Bluestone,J.A. and Matis,L.A.
Identification of Rlk, a novel protein tyrosine kinase with
predominant expression in the T cell lineage
J. Biol. Chem. 270 (4), 1928-1934 (1995)
95130578
PUBMED 7829530
COMMENT On Jan 14, 1995 this sequence version replaced gi:529439.
Original source text: Mus cckii (clone K9a) tetus thymus cDNA to
mRNA.
FEATURES
source
Location/Qualifiers
1..2221
/organism="Mus cckii"
/mol_type="mRNA"
/db_xref="taxon:10098"
/clone="K9a"
/tissue_type="thymus"
/dev_stage="tetus"
1..2221
/gene="RLK"
1..2221
/gene="RLK"
/product="resting lymphocyte kinase"
/notes="predominant expression in T cell lineage; protein
tyrosine kinase"
evidence=experimental
1..48
/gene="RLK"
49..1632
/gene="RLK"
/standard_name="protein tyrosine kinase"
/codon_start=1
/evidence=experimental
/product="resting lymphocyte kinase"
/protein_id="AAA67039.1"
/db_xref="GI:623443"

```

```

/translation="MILSSYSSFQSVLCCCRCSVQKROVTRQISSLREBELSEKHS
ORQRPWFAKMGKTQSNRGVQPSRKKEPLPQPPDERIQVAKLYDFLPREPGLA
LKRAEVLILERCOPHWKARDFGNEGLIFSNVTENRNLANLEIYWHKNTIRNOT
ERLQKAKAGATIVDRSHLGSYITISVTRARRHTQSSIKHYQIKNDGQWYITER
HLFSPVELIOYHOYNAAGLISRLRYPIGLLGSCLPATSGFSYKEMIDPSLAFVKE
IGSQFGVHLGSEWRAHIPVAIKAINEGMSEEDFIEAKVMKLSHSLVLYGVC
OQKPLYITVTFMGCLLDYLRERGOQLKALLSMCODICEGMAYLERSCVIHRDLA
ARNCVSSACVKSIDFGMARYVLDDEVYISSGAKFPVKMCPPEVPHFNKYSKSDVM
SFGLVMEVTEGKMPFENKSNLVQVVEAISQGFRLYRPHLAPMTIYRVVMSCHWSPK
GRPTFAELLQVLTEIAETW"
49..1629
/gene="RLK"
/product="resting lymphocyte kinase"
/standard_name="protein tyrosine kinase"
/evidence=experimental
1633..2221
/gene="RLK"
2193..2198
/gene="RLK"
/evidence=experimental
2221
/gene="RLK"
/evidence=experimental

mat_peptide
ORIGIN
Query Match 1.0%; Score 24; DB 10; Length 2221;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 GAGCAGCGGCAACGACGACGACGAG 58
|||||
Db 110 GAGCAGCGGCAACGACGACGACGAG 87

RESULT 65
MUSPTKRL18/c
LOCUS Mus musculus PTK-RL-18 mRNA for protein tyrosine kinase, complete
DEFINITION cds.
ACCESSION D43963
VERSION D43963.1 GI:604883
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Higashitsuji,H., Nonoguchi,K., Arii,S., Furutani,M., Kaneko,Y.,
Nakayama,H. and Fujita,J.
Enhanced expression of multiple protein tyrosine kinases in the
mouse regenerating liver : Isolation of PTK-RL-18, a novel
cytoplasmic tyrosine kinase gene of Tec PTK family
Unpublished
2 (bases 1 to 2342)
Higashitsuji,H.
Direct Submission
Submitted (19-DEC-1994) Hiroaki Higashitsuji, Faculty of Medicine,
Kyoto University, Clinical Molecular Biology, Shogoin Kawahara-cho
54, Sakyo-ku, Kyoto, Kyoto 606-8507, Japan
(E-mail:higashitsuji@viru.kyoto-u.ac.jp, Tel:81-75-751-3753,
Fax:81-75-751-3750)
FEATURES
source
Location/Qualifiers
1..2342
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57/BL6"
/db_xref="taxon:10090"
/clone="ptk-18"
/sex="male"
/tissue_type="liver"
/clone_lib="Uni-Zap XR"
/dev_stage="adult"
1..2342
/gene

```

```
/gene="PTK-RL-18"
168..1751
/gene="PTK-RL-18"
/codon_start=1
/evidence=experimental
/product="protein tyrosine kinase"
/protein_id="BAA07900.1"
/db_xref="GI:604884"
/translation="MISFSDSSFOSVLCVCCRCVSKROVRTQISLSEELSEKHS
QRQPFALMGKTQNSREGVQVSKRKLPLPQEPDERIQVKALEDLPRPQNL
LRABEYLILERCDFHWKARDRFGNEGLIPSNVTENLANLEIYWHKNTROT
ERLRQAKGAGFIVRSRHLSYTSVFFTRARRHTQSSIKHYQIKKNDGQWYITER
HLFSPVELIQHYNAAGLISRLRYPIGLISCLPATSGYKWEIDPSLTFVKE
IGSQGFVHLGWEAHPVAIKAINEGSSEEDFIEBAKVMKLSHSLRYVLYGVCI
QOKPLVITVFMENGCLDYLREKQLOKALLSMQDICEGMAYLERSYIHRDLA
ARNCLVSACVVKISDFGMARYVLDDEYIISSGAKFPVKCPPEVHFENKYSKSDVW
SFGVLMEVFTGKMPENKSNLQVVAISQGFRLYRPLAPMTIYSVMYSCWHESPK
GRPTFAELLQVLTETATW"
polya_signal
2317..2322
/gene="PTK-RL-18"

ORIGIN
Query Match 1.0%; Score 24; DB 10; Length 2342;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 GAGCAGCGGCAACAGCAGCAGCAG 58
|||||
Db 229 GAGCAGCGGCAACAGCAGCAGCAG 206
|||||

RESULT 66
AC013623_2
WPCOMMENT
Sequence split into 5 fragments LOCUS AC013623 Accession AC013623
Fragment Name Begin End
AC013623_0 1 110000
AC013623_1 100001 210000
AC013623_2 200001 310000
AC013623_3 300001 410000
AC013623_4 400001 502225
Continuation (3 of 5) of AC013623 Mus musculus clone RP23-651

Query Match 1.0%; Score 24; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 GAGCAGCGGCAACAGCAGCAGCAG 58
|||||
Db 105028 GAGCAGCGGCAACAGCAGCAGCAG 105051
|||||

RESULT 67
AC013623_3
WPCOMMENT
Sequence split into 5 fragments LOCUS AC013623 Accession AC013623
Fragment Name Begin End
AC013623_0 1 110000
AC013623_1 100001 210000
AC013623_2 200001 310000
AC013623_3 300001 410000
AC013623_4 400001 502225
Continuation (4 of 5) of AC013623 Mus musculus clone RP23-651

Query Match 1.0%; Score 24; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 GAGCAGCGGCAACAGCAGCAGCAG 58
|||||
Db 5028 GAGCAGCGGCAACAGCAGCAGCAG 5051
|||||

RESULT 68
CR382132_00
WPCOMMENT
Sequence split into 40 fragments LOCUS CR382132 Accession CR382132
Fragment Name Begin End
CR382132_00 1 110000
CR382132_01 100001 210000
CR382132_02 200001 310000
CR382132_03 300001 410000
CR382132_04 400001 510000
CR382132_05 500001 610000
CR382132_06 600001 710000
CR382132_07 700001 810000
CR382132_08 800001 910000
CR382132_09 900001 1010000
CR382132_10 100001 1110000
CR382132_11 1100001 1210000
CR382132_12 1200001 1310000
CR382132_13 1300001 1410000
CR382132_14 1400001 1510000
CR382132_15 1500001 1610000
CR382132_16 1600001 1710000
CR382132_17 1700001 1810000
CR382132_18 1800001 1910000
CR382132_19 1900001 2010000
CR382132_20 2000001 2110000
CR382132_21 2100001 2210000
CR382132_22 2200001 2310000
CR382132_23 2300001 2410000
CR382132_24 2400001 2510000
CR382132_25 2500001 2610000
CR382132_26 2600001 2710000
CR382132_27 2700001 2810000
CR382132_28 2800001 2910000
CR382132_29 2900001 3010000
CR382132_30 3000001 3110000
CR382132_31 3100001 3210000
CR382132_32 3200001 3310000
CR382132_33 3300001 3410000
CR382132_34 3400001 3510000
CR382132_35 3500001 3610000
CR382132_36 3600001 3710000
CR382132_37 3700001 3810000
CR382132_38 3800001 3910000
CR382132_39 3900001 4003362
LOCUS CR382132 4003362 bp DNA linear PLN 01-SEP-2004
DEFINITION Yarrowia lipolytica chromosome F of strain CLIB99 of Yarrowia
lipolytica.
ACCESSION CR382132
VERSION CR382132.1 GI:49650673
KEYWORDS genomic DNA.
SOURCE Yarrowia lipolytica CLIB99
ORGANISM Yarrowia lipolytica CLIB99
Saccharomycetales; Dipodascaceae; Yarrowia.
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
REFERENCE 1 (bases 1 to 4003362)
AUTHORS Dujon,B., Sherman,D., Fischer,G., Durrans,P., Casaregola,S.,
Lafontaine,I., De Montigny,J., Marck,C., Neuvéglise,C., Talla,E.,
Goffard,N., Frangeul,L., Aigle,M., Anthouard,V., Babour,A.,
Barbe,V., Barnay,S., Blanchin,S., Beckerich,J.M., Beyne,E.,
Bleykasten,C., Boissrame,A., Boyer,J., Cattolico,L.,
Confanier,I., F., De Daruvar,A., Despons,L., Fabre,E., Fairhead,C.,
Ferry-Dumazet,H., Groppi,A., Hantraye,F., Hennequin,C.,
Jauniaux,N., Joyet,P., Kachouri,R., Kerrest,A., Koszul,R.,
Lemaire,M., Lesur,I., Ma,L., Muller,H., Nicaud,J.M., Nikolski,M.,
Ozias S., Ozier-Kalogeropoulos,O., Pellenz,S., Potier,S.,
Richard,G.F., Sraub,M.L., Suleau,A., Swennen,D., Tekaiia,F.,
Wesolowski-Louvel,M., Westhof,E., Wirth,B., Zeniou-Meyer,M.,
Zivanovic,I., Bolotin-Fukuhara,M., Thierry,A., Boucher,C.,
Caudron,B., Scarpelli,C., Gaillardin,C., Weissenbach,J., Wincker,P.
and Souciet J.L.
CONSRMT Genome evolution in yeasts
TITLE
```



JOURNAL  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Nature 430 (6995), 35-44 (2004)  
15229592  
2 (bases 1 to 4003362)  
Genoscope.

Direct Submission  
Submitted (01-JUL-2004) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr -

Web : www.genoscope.cns.fr)  
This sequence is unfinished. Sequencing was performed at Genoscope  
and annotations were obtained by the genolevures Consortium.

FEATURES  
source

1. 4003362  
/organism="varrowia lipolytica CLIB99"  
/mol\_type="genomic DNA"  
/strain="CLIB99"  
/db\_xref="taxon:284591"  
/chromosome="F"  
/note="Genoscope sequence ID : YALIOFCHR"

complement (13437..14069)

/locus tag="YALIOF00110g"

complement (13437..14069)

/locus tag="YALIOF00110g"

/note="unnamed protein product; no similarity, possibly  
noncoding, hypothetical start"

/codon\_start=1

/protein\_id="CAG77606.1"

/db\_xref="GI:49650674"

/db\_xref="UniProt/TREMBL:Q6C3F8"

/translation="MTNLFNFKRWLRSGARSAKTKWKKPGVSRSLRSLFWTRAT  
CTSPVQEQSPVPEILIGPGTLLGDSAPITLSDPKITDKKGGVNGNSLQSPQEQ  
DCVPSVDPVTRPKSRSCIPILSDLOLRELOSERPDSILINKRKIRKSGQGCDG  
IEVLRHPFFAASGARNLSDCSGTFQCFPEIIEAVSYSLRKNENIPA"

complement (14272..15570)

/locus tag="YALIOF00132g"

complement (14272..15570)

/locus tag="YALIOF00132g"

/note="unnamed protein product; weakly similar to  
tr|Q12161 Saccharomyces cerevisiae YOL054W, hypothetical  
start"

/codon\_start=1

/protein\_id="CAG77607.1"

/db\_xref="GI:49650675"

/db\_xref="UniProt/TREMBL:Q6C3F7"

/translation="MSTDIRFEAHGSDRVAGLLCDVGSDLCAICQSIMFVPFVLD  
CGHNYCPKQWVNNSCPCEKTKIRKAPTHSLTLKSLAESITWKMIGVDPISLEKX  
PAQRKEQVPYAKOIKKRVFPAPVAESSVHLDEEDGVORCSQHWELREGEDQTN  
CGVWVDSGREAAESINVDDESDSDSQPGEDIIDYDDGVDDVDDVDDGGVTC  
DVTILSDSQASSPAPSTVQINDSLDSDDEOFLPKABSTEVTLDDSDDEQLQPR  
RKHKIMSDSDHPVAVSRKSKRKIPSDSDDDNQOQSSKSKATITFDSEDDDERP  
RVRLKRRDPAVVLSDEREESSESDSDYSDTDSNGITGVYTGLPFSRRRRGGDG  
DENSDSDSDSDNDLGGFVVDREADGSDDEAYSEY"

complement (15838..18501)

/locus tag="YALIOF00154g"

complement (15838..18501)

/locus tag="YALIOF00154g"

/note="unnamed protein product; similar to sp|O13396  
Neurospora crassa MSH-2 DNA mismatch repair protein MSH2,  
start by similarity"

/codon\_start=1

/protein\_id="CAG77608.1"

/db\_xref="GI:49650676"

/db\_xref="UniProt/TREMBL:Q6C3F6"

/translation="MSRPFNDITEPSLTAFAVKGLEEREPDNGGTIRVFERPDG  
YVFGVDARVAVRYRLTKTDLKQEVNITAGSQNFIRDALNGLKVEIYG  
KVDGKMTSMGSPGNLSQVEDLMSGQLNTNPAIAVKTGDTVGLVLDNNHVLGV  
SEFEDNECHNLSLLIQDVKECITSDKRVIAERAGVSRITDAKSNFNAVEEN  
LDNLAKLPQTSPELSKNGALSLACLIKYLSTDSANHGAFKTKHTLSQYMKLD  
ASAKALHMQTSVKDSTKSSLYGLLVCKTATGSRTLAQWQKPMQKQIEERHID  
VEFTSDLESIRQLNSTIPDLNRLTKFMQASLEDVVRVQWATLPHIASGLR  
AAOSELEFTTOLDISTITGLQFELVETDLSIDSHFMINPWERGLANDTKA  
RLACQDRKMDIPASVDELGMEMDKLKFENHHVHGFSRLTRDASCPLRGSAVL  
LATLKAGIIFTNELRSLNFTDLSQYKKIOARLAKETIIEACSCPLRGSAVL  
GOLDVLTSPASVAIRNVIRTVVDSDDRKILTLTASRHPCLEAQDTFTPNVHLGQDS

KKFLVITGPNWGKSTFIQVGVIVLMNOIGCFVPCDRAEISIFDCILARVAGDSQL  
KGLSTFMSLETSAILKATSKSLIIDELGRGTSTYDGLAWAISEHIVQNCFS  
MFATHFELTELAKHPDRVDNLHVAARHVGSSDDITLLYKRVFVGSYSYGHVAEV  
VFPKTVVNAKRAQELDDVNSGTQGGKYASEDLVAGNKLKLEILTEWKSQIKGDEV  
DGAQLKTVVDKYTKTQMBQNVFINDALASL"

complement (19703..20934)

/locus tag="YALIOF00176g"

complement (19703..20934)

/locus tag="YALIOF00176g"

/note="unnamed protein product; similar to DEHA0G03828g

Debaryomyces hansenii IPF 785.1 and sp|P20107

Saccharomyces cerevisiae YMR243c ZRC1 zinc- and cadmium

resistance protein, start by similarity"

/codon\_start=1

/protein\_id="CAG77609.1"

/db\_xref="GI:49650677"

/db\_xref="UniProt/TREMBL:Q6C3F5"

/translation="MSLTREIRMIALLIIDTCFFLLEAIVGVAVHSLALVADSPHML  
NDVFSLIILAWAVRAVSRGADSKYTGQWAEILGALANAVFLALCILLIEAIOQR  
LFEPQIITNPKLIAVVTAGLCNSITVGLLFHEHGHAGSHGHHDHHDHARDEEA  
VDTVLASQFTAPTEQTSLLQHPSTSHRRSISNDSSEHATPHAKKNEQKQVSLNMQ  
GVFLHVMGDALGNIGVIATAFIMKTYSWKYADPVISLVTIIFSSALPLCRSTS  
SILIQAVPQINAEVDNEIVALDQGBELHDLHIWLKEDTFAVILVGVADSPSEFM  
TLSNDIKKIFHEGINSVTIQPEFNATGTTTPDKHQYHVSVGSLRSANSNGLAQ"

complement (19703..20796)

/locus tag="YALIOF00176g"

number=2

complement (20797..20852)

/locus tag="YALIOF00176g"

number=1

complement (20853..20934)

/locus tag="YALIOF00176g"

number=1

22939..25541

/locus tag="YALIOF00198g"

join (22939..22994,23108..25541)

/locus tag="YALIOF00198g"

/note="unnamed protein product; similar to tr|Q08951

Saccharomyces cerevisiae ORF YPL195W Alpha/gamma adaptin,

start by similarity"

/codon\_start=1

/protein\_id="CAG77610.1"

/db\_xref="GI:49650678"

/db\_xref="UniProt/TREMBL:Q6C3F4"

/translation="MSLTTRDDVRNRLRPFPIESTLTDLIKIRAAKDLPAKERFL  
QAVSDCHQEARNDLTAKTLALILKALAYLEMLGYDMSWASHVLEVMSPKQQRV  
YLAATOSYRDDDDVLMATLNLKRLDTSSEFVSLSAIAITVPSLAQDVHPEIV  
KMLSHSKYIRKXAVLAMYKIFLOYPEALRTSFLRDLRDDSDPSVVSATVNVICEL  
AKHKSNNYPLAPOLYQLLTSSNNMMIKILFSSLAPIEPRLPKPLQIMTLIQ  
STGSLQYECINCIIVSGMLGEDDHVAGVCVCSFVNQDQNLKVLGALGKI  
KVHPLTLVGLQVILECLKENADSTIRERALELANDLASEHNVTIVNLISQOLTH  
ASISYILDMCCRDYSLISDPFNVLKALVDFPDLPARIGETLRLDCMRVPDM  
REBITGVCVEIFTDGVTNRNPGVLPVLPVLCSEYPMHQKDLDFCLNSHVRENS  
GSSNNAVDLAVCVICVKGKFCGEGIIWSSPNELVYKTTVTVIKELSYLASL  
QDFATQFSLFKVIESIESEERGVPVLTLEVPSPFAYELNVPADTKRIPLDLLEP  
PELEQTLEVFSDIEDEYVEEFTQPQASTVSGVQKESLIDARKERLEKQKODFFI

Query Match 1.0%; Score 24; DB 8; Length 110000;

Best Local Similarity 100.0%; Pred. No. 0.86;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 AGCAGCGGCAACAGCAGCAGCAGCAGC 59

Db 93631 AGCAGCGGCAACAGCAGCAGCAGCAGC 93654

RESULT 69

BX663616/c

LOCUS BX663616

DEFINITION Zebrafish DNA sequence from clone DKEY-287H22 in linkage group 7,

complete sequence.

ACCESSION BX663616

VERSION BX663616.13 GI:47522521

KEYWORDS	HTG.	ACCESSION	AL732455
SOURCE	Danio rerio (zebrafish)	VERSION	AL732455.6 GI:21953067
ORGANISM	Danio rerio	KEYWORDS	HTG; zinc finger.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.	SOURCE	Danio rerio (zebrafish)
AUTHORS	1 (bases 1 to 150010)	ORGANISM	Danio rerio
TITLE	Auger,K.	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
JOURNAL	Direct Submission	1 (bases 1 to 192578)	1 (bases 1 to 192578)
COMMENT	Submitted (19-MAY-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On May 19, 2004 this sequence version replaced gi:47109759.	SKUCE,C.	SKUCE,C.
	----- Genome Center	Direct Submission	Direct Submission
	Center code: SC	Submitted (20-JUN-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Jul 24, 2002 this sequence version replaced gi:21912710.	Submitted (20-JUN-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Jul 24, 2002 this sequence version replaced gi:21912710.
	Contact: zfish-help@sanger.ac.uk	During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.	During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
	-----	This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.	This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
	During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.	This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.	This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
	This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.	The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep	The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep
	is from a Zebrafish BAC library	VECTOR: pindigoBAC-5.	VECTOR: pindigoBAC-5.
FEATURES	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers
source	1..150010	source	1..192578
	/organism="Danio rerio"		/organism="Danio rerio"
	/mol_type="genomic DNA"		/mol_type="genomic DNA"
	/db_xref="taxon:7955"		/db_xref="taxon:7955"
	/clone_lib="Daniokey"		/clone_lib="CH211-225J17"
ORIGIN			
Query Match	1.0%; Score 24; DB 5; Length 150010;	repeat_region	1..423
Best Local Similarity	100.0%; Pred. No. 0.87; Mismatches 0; Indels 0; Gaps 0;	repeat_region	/note="Dr000087 repeat: matches 255. .674 of consensus"
Matches 24; Conservative 0;		repeat_region	complement(556..661)
		repeat_region	/note="drr831 repeat: matches 1. .105 of consensus"
QY	1993 GAAGAGGACAGATCAATATTCAC 2016	repeat_region	654..1032
		repeat_region	/note="Dr000087 repeat: matches 957. .1326 of consensus"
Db	29978 GAAGAGGACAGATCAATATTCAC 29955	repeat_region	1033..1089
		repeat_region	/note="19.0 copies 3 mer TTTTAA 28% conserved"
RESULT 70	192578 bp DNA linear VRT 29-JUL-2003	repeat_region	1114..1127
AL732455		repeat_region	/note="2.3 copies 6 mer TTTTAA 28% conserved"
LOCUS		repeat_region	1145..1207
DEFINITION	Zebrafish DNA sequence from clone CH211-225J17 in linkage group 7	repeat_region	/note="CRI-1_DR repeat: matches 4919. .4981 of consensus"
	Contains part of a novel gene for a zinc finger protein, complete sequence.	repeat_region	1203..1268

```
repeat_region /note="ACROBAT1 repeat: matches 776. .842 of consensus"
repeat_region complement(1214. .1275)
repeat_region /note="Dr000340 repeat: matches 135. .196 of consensus"
repeat_region complement(1272. .1327)
repeat_region /note="Dr000403 repeat: matches 1. .66 of consensus"
repeat_region complement(1369. .1617)
repeat_region /note="Dr000262 repeat: matches 5. .284 of consensus"
repeat_region 1620. .1631
repeat_region /note="12.0 copies 1 mer T 24% conserved"
repeat_region 1659. .1731
repeat_region /note="Dr000067 repeat: matches 207. .288 of consensus"
repeat_region 1736. .1747
repeat_region /note="2.0 copies 6 mer TTAGTA 24% conserved"
repeat_region 1774. .1795
repeat_region /note="2.8 copies 8 mer TTTAAGA 26% conserved"
repeat_region 1998. .2008
repeat_region /note="2.8 copies 4 mer AGAT 22% conserved"
repeat_region 2017. .2031
repeat_region /note="3.8 copies 4 mer CAGG 30% conserved"
repeat_region 2032. .2102
repeat_region /note="HAT1_DR repeat: matches 575. .647 of consensus"
repeat_region 2103. .2200
repeat_region /note="HAT1_DR repeat: matches 555. .660 of consensus"
repeat_region 2135. .2211
repeat_region /note="Dr000309 repeat: matches 2. .78 of consensus"
repeat_region 2397. .2412
repeat_region /note="8.0 copies 2 mer CT 23% conserved"
repeat_region 2471. .2481
repeat_region /note="2.2 copies 5 mer AATAA 22% conserved"
repeat_region 2930. .2943
repeat_region /note="7.0 copies 2 mer TC 28% conserved"
repeat_region 3063. .3074
repeat_region /note="2.4 copies 5 mer ATATG 24% conserved"
repeat_region 3280. .3407
repeat_region /note="HAT1_DR repeat: matches 521. .659 of consensus"
repeat_region 3371. .3442
repeat_region /note="Dr000309 repeat: matches 2. .69 of consensus"
repeat_region 3408. .3448
repeat_region /note="HAT1_DR repeat: matches 577. .617 of consensus"
repeat_region 3456. .3467
repeat_region /note="2.0 copies 6 mer AACAG 24% conserved"
repeat_region 3629. .3641
repeat_region /note="13.0 copies 1 mer T 26% conserved"
repeat_region 3687. .3696
repeat_region /note="2.5 copies 4 mer ATAT 20% conserved"
repeat_region 3807. .3819
repeat_region /note="2.2 copies 6 mer TGATTC 26% conserved"
repeat_region 3914. .3947
repeat_region /note="8.5 copies 4 mer TAAA 50% conserved"
repeat_region 4157. .4273
repeat_region /note="58.5 copies 2 mer CA 198% conserved"
repeat_region 4174. .4195
repeat_region /note="3.1 copies 7 mer ACACTAA 23% conserved"
repeat_region 4274. .4346
repeat_region /note="Dr000236 repeat: matches 358. .433 of consensus"
repeat_region 4642. .4653
repeat_region /note="2.0 copies 6 mer GTTTC 24% conserved"
repeat_region 4726. .4904
repeat_region /note="DNA11TA1_DR repeat: matches 1. .186 of consensus"
repeat_region complement(4960. .5090)
repeat_region /note="ACROBAT1 repeat: matches 751. .876 of consensus"
repeat_region complement(5134. .5259)
repeat_region /note="DNA11TA1_DR repeat: matches 1. .130 of consensus"
repeat_region 5413. .5423
repeat_region /note="2.2 copies 5 mer TACAT 22% conserved"
repeat_region complement(5598. .5681)
repeat_region /note="Dr000009 repeat: matches 770. .857 of consensus"
repeat_region 5803. .5824
repeat_region /note="22.0 copies 1 mer T 44% conserved"
repeat_region 6147. .6167
repeat_region /note="2.3 copies 9 mer ACAGAATGG 35% conserved"
repeat_region 6224. .6233
repeat_region /note="2.5 copies 4 mer ACAA 20% conserved"
```

```
repeat_region 6238. .6253
repeat_region /note="16.0 copies 1 mer A 32% conserved"
repeat_region 6626. .6635
repeat_region /note="2.5 copies 4 mer GAAT 20% conserved"
repeat_region 6738. .6760
repeat_region /note="7.7 copies 3 mer TCC 39% conserved"
repeat_region complement(7233. .7249)
repeat_region /note="TDR19 repeat: matches 198. .303 of consensus"
repeat_region 7250. .7336
repeat_region /note="Dr000364 repeat: matches 4. .90 of consensus"
repeat_region 7336. .8045
repeat_region /note="Dr000209 repeat: matches 1. .678 of consensus"
repeat_region 8045. .8138
repeat_region /note="Dr000364 repeat: matches 86. .180 of consensus"
repeat_region 8189. .8203
repeat_region /note="15.0 copies 1 mer T 30% conserved"
repeat_region 8224. .8245
repeat_region /note="5.5 copies 4 mer TCAA 28% conserved"
repeat_region 8507. .8536
repeat_region /note="10.0 copies 3 mer TAC 51% conserved"
repeat_region 8542. .8609
repeat_region /note="22.7 copies 3 mer ATA 118% conserved"
repeat_region 8686. .8703
repeat_region /note="4.5 copies 4 mer CATT 27% conserved"
repeat_region 8705. .8714
repeat_region /note="2.5 copies 4 mer TCGT 20% conserved"
repeat_region 8712. .8769
repeat_region /note="14.5 copies 4 mer TTCC 91% conserved"
repeat_region 8771. .8794
repeat_region /note="6.0 copies 4 mer TTCA 48% conserved"
repeat_region 8828. .8837
repeat_region /note="2.5 copies 4 mer TTAC 20% conserved"
repeat_region 8909. .8919
repeat_region /note="2.8 copies 4 mer TATC 22% conserved"
repeat_region complement(8972. .9131)
repeat_region /note="HEI_DR1 repeat: matches 41. .221 of consensus"
repeat_region 9314. .9344
repeat_region /note="5.2 copies 6 mer ACTATG 30% conserved"
repeat_region 9420. .9432
repeat_region /note="6.5 copies 2 mer AG 26% conserved"
repeat_region 9430. .9441
repeat_region /note="2.4 copies 5 mer AGAA 24% conserved"
repeat_region 9447. .9466
repeat_region /note="2.2 copies 9 mer TTACAGAT 31% conserved"
repeat_region 10208. .10222
repeat_region /note="3.8 copies 4 mer AGGA 21% conserved"
repeat_region 10601. .10611
repeat_region /note="2.8 copies 4 mer CTTC 22% conserved"
repeat_region 10867. .10885
repeat_region /note="3.2 copies 6 mer ACTGTA 29% conserved"
```

Query Match 1.0%; Score 24; DB 5; Length 192578;  
Best Local Similarity 100.0%; Pred. No. 0.87;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1993 GAACAGCAGACATCAATATTCCA 2016  
Db 6479 GAACAGCAGACATCAATATTCCA 6502

RESULT 71  
AC099698

LOCUS AC099698 222930 bp DNA linear HTG 30-AUG-2003  
DEFINITION Mus musculus chromosome 5 clone RP23-41509 map 5, \*\*\* SEQUENCING IN  
PROGRESS \*\*\*, 10 unordered pieces.

ACCESSION AC099698.6 GI:34365951  
VERSION HTG; HTGS PHASE1; HTGS FULLTOP; HTGS\_ACTIVEFIN.  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus

ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 222930)

Birren,B., Nusbaum,C. and Lander,E.  
 Mus musculus chromosome 5, clone RP23-41509  
 Unpublished  
 2 (bases 1 to 222930)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,  
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,  
 Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A.,  
 Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,  
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,  
 Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K.,  
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,  
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
 McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,  
 Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,  
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,  
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
 Roman,J., Rosetti,M., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (18-NOV-2001) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 222930)  
 Birren,B., Nusbaum,C. and Lander,E.  
 Direct Submission  
 Submitted (30-AUG-2003) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Aug 30, 2003 this sequence version replaced gi:31581768.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 ----- Genome Center / MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence.submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L17169  
 Center clone name: 415\_O\_9  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 10 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence.  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 35963: contig of 35963 bp in length  
 \* 35964: gap of 100 bp  
 \* 36064: contig of 4705 bp in length  
 \* 40769: gap of 100 bp  
 \* 40868: contig of 3639 bp in length  
 \* 44507: gap of 100 bp  
 \* 44508: contig of 5291 bp in length  
 \* 44508: contig of 5291 bp in length  
 \* 49898: gap of 100 bp  
 \* 49899: contig of 4328 bp in length  
 \* 54327: gap of 100 bp  
 \* 54427: contig of 81601 bp in length  
 \* 136027: gap of 100 bp  
 \* 136128: contig of 17672 bp in length  
 \* 153800: gap of 100 bp  
 \* 153800: contig of 37218 bp in length  
 \* 191118: gap of 100 bp  
 \* 191218: contig of 3759 bp in length  
 \* 194977: gap of 100 bp

195077 222930: contig of 27854 bp in length.  
 Location/Qualifiers  
 1. 222930  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosomes="5"  
 /map="5"  
 /clone="RP23-41509"  
 /clone\_lib="RP23 Female Mouse BAC"  
 ORIGIN  
 Query Match 1.0%; Score 24; DB 2; Length 222930;  
 Best Local Similarity 100.0%; Pred. NO. 0.87;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 35 GAGCAGCGGCACAGCAGCAGCAG 58  
 DB 149599 GAGCAGCGGCACAGCAGCAGCAG 149622  
 RESULT 72  
 AC102441/c  
 LOCUS  
 DEFINITION  
 Mus musculus chromosome 18, clone RP24-160112, complete sequence.  
 AC102441  
 VERSION  
 AC102441.6 GI:32526991  
 KEYWORDS  
 HTG.  
 SOURCE  
 Mus musculus (house mouse)  
 ORGANISM  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 Birren,B., Nusbaum,C. and Lander,E.  
 Mus musculus chromosome 18, clone RP24-160112  
 Unpublished  
 2 (bases 1 to 224556)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,  
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,  
 Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A.,  
 Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,  
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,  
 Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K.,  
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,  
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
 McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,  
 Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,  
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,  
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
 Roman,J., Rosetti,M., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 224556)  
 Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,  
 Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,  
 Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,  
 Collymore,A., Cooke,A., Cooke,P., Corum,B., DeArellano,K.,  
 Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,  
 Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,  
 Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,  
 Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,  
 Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,

Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (12-JUN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 224556)  
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (11-JUL-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

5 (bases 1 to 224556)  
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (30-SEP-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jul 11, 2003 this sequence version replaced gi:31621396.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WBIR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center Project name: L18799  
Center Clone name: 160\_I\_12

FEATURES  
Source  
----- Location/Qualifiers  
1..224556  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="18"  
/map="18"  
/clone="RP24-160I12"  
/clone\_lib="RPCI-24 Male Mouse BAC"  
1..4389  
/note="wgs end extension  
clone end:SP6"  
complement(1..55)  
/note="<30 qual SNGL region"  
complement(156..186)  
/rpt\_family="B2\_Mm2"  
187..229  
/rpt\_family="(TTTA)n"  
complement(230..390)  
/rpt\_family="B2\_Mm2"  
392..983  
/rpt\_family="RMER19B"  
1012..1306  
/rpt\_family="RMER19B"  
complement(1509..1652)  
/rpt\_family="B1\_MM"  
complement(1658..1857)  
/rpt\_family="B2\_Mm1"  
complement(1868..1998)  
/rpt\_family="HAL1"  
complement(2049..2185)  
/rpt\_family="B1\_MM"  
complement(2225..2408)  
/rpt\_family="B2\_Mm1"  
complement(2436..2628)  
/rpt\_family="B3"  
complement(2698..2858)  
/rpt\_family="ID\_B1"  
3272..3276  
/note="<30 qual SNGL region"  
3325..3329  
/note="<30 qual SNGL region"  
3338..3343  
/note="<30 qual SNGL region"  
3354..3358  
/note="<30 qual SNGL region"  
3534..3586  
/rpt\_family="AT\_rich"  
3651..3690  
/rpt\_family="AT\_rich"  
3719..3747  
/rpt\_family="(TA)n"  
3995..4036  
/rpt\_family="AT\_rich"  
4040..4270  
/rpt\_family="B3"  
4390..4393  
/note="clone boundary  
clone end:SP6  
site:MboI"  
7019..7088  
/rpt\_family="ID5"  
7303..7323  
/rpt\_family="AT\_rich"  
7481..7504  
/rpt\_family="AT\_rich"  
7961..8042  
/rpt\_family="ID4"  
8150..8331  
/rpt\_family="B2\_Mm2"  
8595..8623  
/rpt\_family="AT\_rich"

```

repeat_region complement(8747..8913)
                /rpt_family="L1ME3B"
repeat_region complement(9211..9354)
                /rpt_family="B1_MM"

Query Match      1.0%; Score 24; DB 10; Length 224556;
Best Local Similarity 100.0%; Pred. No. 0.87;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 AGCAGCGCGCACAGCAGCAGCAGC 59
    |||||
Db 179363 AGCAGCGCGCACAGCAGCAGCAGC 179340

RESULT 73
AC090126/c
LOCUS
DEFINITION Mus musculus chromosome 5 clone RP23-6518 map 5, *** SEQUENCING IN
PROGRESS ***, 11 unordered pieces.
AC090126
AC090126.4 GI:30103086
VERSION HTGS HTGS PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Birren,B., Nusbaum,C. and Lander,E.
1 (bases 1 to 289405)
Mus musculus chromosome 5, clone RP23-6518
Unpublished
2 (bases 1 to 289405)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguslavsky,L., Bouckgalter,B., Brown,A.,
Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collimore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Headford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrum,J., Meneus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (17-FEB-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 289405)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavsky,L., Bouckgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collimore,A., Cooke,P., Cooke,K., Dorris,B., DeArellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
Macdonald,P., Major,J., Manning,J., Mienga,V., Murphy,T., Naylor,J.,
Nguyen,C., Meneus,L., Mihova,T., Meneus,L., Mienga,V., McCarthy,M.,
O'Donnell,P., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,

```

# TITLES JOURNAL

## COMMENT

Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,  
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,  
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (25-APR-2003) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 25, 2003 this sequence version replaced gi:30018220.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information  
Center project name: L11591  
Center clone name: 65\_1\_8

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 11 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 28999: contig of 28999 bp in length  
\* 29000: gap of 100 bp  
\* 29100: contig of 7707 bp in length  
\* 36806: gap of 100 bp  
\* 36807: gap of 100 bp  
\* 45491: contig of 8585 bp in length  
\* 45492: gap of 100 bp  
\* 45591: contig of 3079 bp in length  
\* 48670: gap of 100 bp  
\* 48671: contig of 3561 bp in length  
\* 52331: gap of 100 bp  
\* 52332: gap of 100 bp  
\* 52431: contig of 15502 bp in length  
\* 52432: gap of 100 bp  
\* 67933: contig of 11546 bp in length  
\* 67934: gap of 100 bp  
\* 79579: contig of 67278 bp in length  
\* 79580: gap of 100 bp  
\* 79680: contig of 145957 bp in length  
\* 146958: gap of 100 bp  
\* 147057: contig of 41535 bp in length  
\* 188592: gap of 100 bp  
\* 188593: contig of 18431 bp in length  
\* 207123: gap of 100 bp  
\* 207124: contig of 82182 bp in length.  
\* 207224: gap of 100 bp  
\* 207242: contig of 82182 bp in length.

Location/Qualifiers  
1. 289405  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="5"  
/map="5"  
/clone\_lib="RP23-6518"  
/clone\_lib="RP23-6518" Female Mouse BAC"

Query Match 1.0%; Score 24; DB 2; Length 289405;  
Best Local Similarity 100.0%; Pred. No. 0.87;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 GAGCAGCGCGCACAGCAGCAGCAGC 58  
|||||

Db 193297 GAGCAGCGCGCACAGCAGCAGCAGC 193274  
|||||

RESULT 74  
CO871636  
LOCUS  
DEFINITION Sequence 127 from Patent WO2004/078783.  
ACCESSION CO871636

QY 35 GAGCAGCGCGCACAGCAGCAGCAGC 58  
|||||

Db 193297 GAGCAGCGCGCACAGCAGCAGCAGC 193274  
|||||

RESULT 74  
CO871636  
LOCUS  
DEFINITION Sequence 127 from Patent WO2004/078783.  
ACCESSION CO871636

QY 35 GAGCAGCGCGCACAGCAGCAGCAGC 58  
|||||

Db 193297 GAGCAGCGCGCACAGCAGCAGCAGC 193274  
|||||

RESULT 74  
CO871636  
LOCUS  
DEFINITION Sequence 127 from Patent WO2004/078783.  
ACCESSION CO871636

QY 35 GAGCAGCGCGCACAGCAGCAGCAGC 58  
|||||

Db 193297 GAGCAGCGCGCACAGCAGCAGCAGC 193274  
|||||

RESULT 74  
CO871636  
LOCUS  
DEFINITION Sequence 127 from Patent WO2004/078783.  
ACCESSION CO871636

QY 35 GAGCAGCGCGCACAGCAGCAGCAGC 58  
|||||

Db 193297 GAGCAGCGCGCACAGCAGCAGCAGC 193274  
|||||

RESULT 74  
CO871636  
LOCUS  
DEFINITION Sequence 127 from Patent WO2004/078783.  
ACCESSION CO871636

QY 35 GAGCAGCGCGCACAGCAGCAGCAGC 58  
|||||

Db 193297 GAGCAGCGCGCACAGCAGCAGCAGC 193274  
|||||

RESULT 74  
CO871636  
LOCUS  
DEFINITION Sequence 127 from Patent WO2004/078783.  
ACCESSION CO871636

QY 35 GAGCAGCGCGCACAGCAGCAGCAGC 58  
|||||

Db 193297 GAGCAGCGCGCACAGCAGCAGCAGC 193274  
|||||

RESULT 74  
CO871636  
LOCUS  
DEFINITION Sequence 127 from Patent WO2004/078783.  
ACCESSION CO871636

QY 35 GAGCAGCGCGCACAGCAGCAGCAGC 58  
|||||

Db 193297 GAGCAGCGCGCACAGCAGCAGCAGC 193274  
|||||

RESULT 74  
CO871636  
LOCUS  
DEFINITION Sequence 127 from Patent WO2004/078783.  
ACCESSION CO871636

QY 35 GAGCAGCGCGCACAGCAGCAGCAGC 58  
|||||

Db 193297 GAGCAGCGCGCACAGCAGCAGCAGC 193274  
|||||

VERSION  
KEYWORDS  
SOURCE  
ORGANISM

CQ871636.1 GI:52745669

synthetic construct  
synthetic construct  
other sequences; artificial sequences.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Murphy, F.J., Sheehan, D.E., Keating, K.E., Hayes, I.E. and Seera, L.E.  
Enzymes involved in apoptosis  
Patent: WO 2004078793-A 127 16-SEP-2004;  
Erix Therapeutics Ltd (IE)

FEATURES  
source

Location/Qualifiers  
1..23  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="qPCR primer"

ORIGIN

Query Match 1.0%; Score 23; DB 6; Length 23;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 CCCGAGAGAGACAAAGCATGGA 128

Db 1 CCCGAGAGAGACAAAGCATGGA 23

RESULT 75

G94809/c  
LOCUS

S209P6230RC12.T0 C3H/HeJ Mus musculus STS genomic, sequence tagged site.

DEFINITION

G94809

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

595 bp DNA linear STS 06-SEP-2002  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Wade, C.  
Polymorphism Structure in the Mouse  
Unpublished (2002)

Contact: Kerstin Lindblad-Toh  
Whitehead Institute for Biomedical Research, Center for Genome  
Research  
320 Charles Street, Cambridge, MA 02141, USA  
Tel: 6172521477

Fax: 6172580903  
Email: kersli@genome.wi.mit.edu  
Primer A: None  
Primer B: None  
STS size: 595

Protocol:

WGS-discovery: Paired-end low-coverage whole genome shotgun reads  
were generated from 129S1/SvimJ, C3H/HeJ, and BALB/cByJ. The WGS  
reads were placed uniquely on the MGSCv3 C57BL/6J assembly and SNP  
detection was carried out by SSAHA-SNP. 225,000 reads were  
annotated  
as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J  
and the strain from which the particular read came. The validation  
rate for these SNPs was estimated at approximately 98%.

FEATURES  
source

Location/Qualifiers  
1..595  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C3H/HeJ"  
/db\_xref="taxon:10090"  
/map="5 22-540 71845760-71846277"  
/clone\_lib="C3H/HeJ"  
<1..->595

STS  
ORIGIN

Query Match 1.0%; Score 23; DB 11; Length 595;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 GAGCAGCGCGCAACAGCAGCAGCA 57

Db 23 GAGCAGCGCGCAACAGCAGCAGCA 1

Search completed: March 25, 2005, 20:35:15  
Job time : 9731 secs



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using ew model

Run on: March 25, 2005, 16:47:45 ; Search time 6778 Seconds  
(without alignments)  
13051.248 Million cell updates/sec

Title: US-09-436-184-3  
Perfect score: 2324  
Sequence: 1 cggaccgtgcaatggccag.....cttgggaactctggagaga 2324

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

- EST.\*
- 1: gb\_est1.\*
  - 2: gb\_est2.\*
  - 3: gb\_est3.\*
  - 4: gb\_est4.\*
  - 5: gb\_est5.\*
  - 6: gb\_est6.\*
  - 7: gb\_est7.\*
  - 8: gb\_est8.\*
  - 9: gb\_est9.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1803	77.6	2126	9	AY405267 Homo sapi
2	766	33.0	1050	4	BM806632 AGENCOURT
3	734	31.6	1085	5	BM459083 BX-EST0070
4	689	29.6	689	4	BM790814 K-EST0070
5	628	27.0	782	7	CK724966 UI-H-ED1
6	603	25.9	918	5	BUI66117 AGENCOURT
7	578	24.9	729	7	CN258120 170005325
8	577	24.8	868	5	BUI75716 AGENCOURT
9	568	24.4	603	5	BP336604 BP336604
10	563	24.2	583	5	BP364957 BP364957
11	556	23.9	859	1	AUI30952 AUI30952
12	543	23.4	651	2	AW996828 QV3-BN004
13	541	23.3	1011	5	EX403636 EX403636
14	538	23.1	589	5	CN258117 170005318
15	537	23.1	581	5	BP266690 BP266690
16	531	22.8	582	5	BP270196 BP270196
17	529	22.8	765	6	CD643948 AGENCOURT
18	517	22.2	517	5	EX479249 DKF2P686L
19	507	21.8	721	4	B1333062 B1333062
20	493	21.2	711	7	CN368832 CN368832
21	491	21.1	868	5	BQ232651 BQ232651
22	476	20.5	520	7	CN368829 CN368829
23	461	19.8	770	7	CN368823 CN368823
24	458	19.7	689	4	BG720220 BG720220

25	455	19.6	455	7	CN274513
26	452	19.4	593	5	BP328261
27	450	19.4	857	5	BQ889709
28	450	19.4	920	5	BQ722578
29	450	19.4	1014	5	BUI90749
30	448	19.3	767	1	AUI32666
31	447	19.2	911	5	BUS02606
32	438	18.8	519	4	BF962729
33	436	18.8	537	2	BE935834
34	430	18.5	695	7	CN368827
35	425	18.3	554	1	AL700393
36	425	18.3	959	5	BUS02544
37	424	18.2	424	1	AL120973
38	424	18.2	606	7	CN481470
39	423	18.2	504	4	BF943222
40	423	18.2	610	7	CN481475
41	420	18.1	790	1	AUI38660
42	418	18.0	433	5	BQ318605
43	417	17.9	1100	5	BQ278505
44	415	17.9	444	4	BF962946
45	409	17.6	570	5	BQ639558
46	402	17.3	740	7	CN368821
47	397	17.1	629	7	CN483754
48	395	17.0	497	5	BQ340475
49	395	17.0	503	5	BQ339559
50	392	16.9	1031	5	BQ223755
51	389	16.7	553	4	BI035330
52	385	16.6	409	2	BE084010
53	383	16.5	930	5	BUI838423
54	382	16.4	468	7	CN368831
55	379	16.3	720	6	CA311352
56	370	15.9	626	7	CN368826
57	329	14.2	1048	2	BE876059
58	328	14.1	574	1	AI267311
59	324	13.9	393	7	CN368818
60	322	13.9	580	7	CN481268
61	321	13.8	527	1	AL694159
62	316	13.6	2023	9	AY405268
63	313	13.5	484	5	EX111251
64	312	13.4	503	7	CF136513
65	311	13.4	682	4	BGI70148
66	308	13.3	719	7	CRG30295
67	308	13.3	737	7	CR559422
68	306	13.2	560	7	CN482623
69	297	12.8	548	7	CN368828
70	295	12.7	509	2	BE018154
71	294	12.7	556	4	BM785890
72	294	12.7	622	1	AV751514
73	294	12.7	862	6	CD512083
74	288	12.4	1010	4	BM462707
75	283	12.2	767	7	CN368820
76	283	12.2	801	5	BUI58234
77	283	12.2	895	4	BI822264
78	278	12.0	692	7	CR769476
79	272	11.7	581	1	AI652752
80	271	11.7	870	5	BX374107
81	270	11.6	468	7	TF9872
82	268	11.5	633	7	CN258115
83	265	11.4	715	4	BG701542
84	264	11.4	273	2	BE004012
85	253	10.9	866	4	BG332543
86	252	10.8	774	4	BI193485
87	249	10.7	316	4	BG977347
88	246	10.6	534	6	CB265658
89	244	10.5	614	4	BM785946
90	243	10.0	534	2	BF154495
91	232	10.0	575	2	BF911695
92	232	10.0	644	4	BI546417
93	232	10.0	789	5	BUS43358
94	232	10.0	793	5	BUS42964
95	231	9.9	927	7	CF619299
96	229	9.9	662	2	AW994191
97	226	9.7	769	4	BGI66775

```
c 98 225 9.7 404 4 BG979400 BG979400 CM4-CN009
99 222 9.6 549 1 AA176259 2P29h01.r
100 221 9.5 504 6 CD675871 f829b12.y

ALIGNMENTS

RESULT 1
LOCUS AY405267 2126 bp DNA linear GSS 16-DEC-2003
DEFINITION Homo sapiens ASPH gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY405267
VERSION AY405267.1 GI:39761241
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
MAMMALIA; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 2126)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 2126)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submision
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..2126
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>2126
/gene="ASPH"
/locus_tag="HCM2167"
gene
ORIGIN
Query Match 77.6%; Score 1803; DB 9; Length 2126;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1803; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 334 GACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCCGAGAGAGCGTCCAGCACACA 393
Db 323 GACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCCGAGAGAGCGTCCAGCACACA 382
Qy 394 CTGAGCCCGAGGAGAGGTTCTGTGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 453
Db 383 CTGAGCCCGAGGAGAGGTTCTGTGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCA 442
Qy 454 AAGAACAAATTCAGTCCCTTCTCCATGAAATGGTACAGCAGAGAGAGAGAGAGAGAGAGAG 513
Db 443 AAGAACAAATTCAGTCCCTTCTCCATGAAATGGTACAGCAGAGAGAGAGAGAGAGAGAGAG 502
Qy 514 ACTTGCAACAGAGAGATGGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 573
Db 503 ACTTGCAACAGAGAGATGGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 562
Qy 574 TGGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 633
Db 563 TGGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 622
```

```
Qy 634 CCGAGCATAGTTTACCAAGTGGAGAGACAGTCTTCAAGAGCTGTATATCAGGATATGGAAG 693
Db 623 CCGAGCATAGTTTACCAAGTGGAGAGACAGTCTTCAAGAGCTGTATATCAGGATATGGAAG 682
Qy 694 AGATGATGCTGAGCAGGAGAAATCCAGATTCAGTGAACCAAGCTAGTAGAGATGAAGAT 753
Db 683 AGATGATGCTGAGCAGGAGAAATCCAGATTCAGTGAACCAAGCTAGTAGAGATGAAGAT 742
Qy 754 TGCACCATGATACAGATGATGTAAACATACCAAGTCTATGAGGAAACAGCAGTATATGAAC 813
Db 743 TGCACCATGATACAGATGATGTAAACATACCAAGTCTATGAGGAAACAGCAGTATATGAAC 802
Qy 814 CTCTAGAAAATGAAGGAGTAGAAAATCAGAGATTAACCTGCTCCCTCCTGAGGATAATCCTG 873
Db 803 CTCTAGAAAATGAAGGAGTAGAAAATCAGAGATTAACCTGCTCCCTCCTGAGGATAATCCTG 862
Qy 874 TAGAAGATTCACAGGTAATTTGTAGAGAGAGTAGACATTTTCCCTGCTGGAGACAGCAGG 933
Db 863 TAGAAGATTCACAGGTAATTTGTAGAGAGAGTAGACATTTTCCCTGCTGGAGAGACAGCAGG 922
Qy 934 AAGTACACCCAGAGAAACAAATAGAAAAACAGATGATCCAGAACAAAAAGCAAAAGTTAAGA 993
Db 923 AAGTACACCCAGAGAAACAAATAGAAAAACAGATGATCCAGAACAAAAAGCAAAAGTTAAGA 982
Qy 994 AAAAGAGCCCTAACTTTTAAATTAATTTGATGAAGACTATTAAGCTGAACCTTGATGCTG 1053
Db 983 AAAAGAGCCCTAACTTTTAAATTAATTTGATGAAGACTATTAAGCTGAACCTTGATGCTG 1042
Qy 1054 CAGAAAAACCTCCGTAAAGGGGAAAAATTCAGGAGACAGTGAATGATTTTAAAGAACTAG 1113
Db 1043 CAGAAAAACCTCCGTAAAGGGGAAAAATTCAGGAGACAGTGAATGATTTTAAAGAACTAG 1102
Qy 1114 TAGCGAAATACCCCTCAGAGTCCACGAGCAAGATATGGGAGGCGCAGTGTGAGGATGATT 1173
Db 1103 TAGCGAAATACCCCTCAGAGTCCACGAGCAAGATATGGGAGGCGCAGTGTGAGGATGATT 1162
Qy 1174 TGCTGAGAGAGAGAGAGATGAGGTGCTAGCTGAGGAGCAATCGAGACCTACCAAGAGG 1233
Db 1163 TGCTGAGAGAGAGAGAGATGAGGTGCTAGCTGAGGAGCAATCGAGACCTACCAAGAGG 1222
Qy 1234 TGCCAGCCCTACCTGATGCTCCCTGAGAGACCTGCTGAGAGCTGAGTTTGAACGCTCGCTCAG 1293
Db 1223 TGCCAGCCCTACCTGATGCTCCCTGAGAGACCTGCTGAGAGCTGAGTTTGAACGCTCGCTCAG 1282
Qy 1294 ACAGGCAACAATTTCTAGGTTCATATGAGAGTTTCCCTGCTTACCTGCGAGAGATTAAGTTC 1353
Db 1283 ACAGGCAACAATTTCTAGGTTCATATGAGAGTTTCCCTGCTTACCTGCGAGAGATTAAGTTC 1342
Qy 1354 AACTATTTCCCAATGATCTCTTAAATAATGACCTTGGCGTGGGATACCTCTTGATAG 1413
Db 1343 AACTATTTCCCAATGATCTCTTAAATAATGACCTTGGCGTGGGATACCTCTTGATAG 1402
Qy 1414 GAGATATGACATGCAAGAGAAAGTTTATGAAGAGGTGCTGAGTGTGACACCTAATGATG 1473
Db 1403 GAGATATGACATGCAAGAGAAAGTTTATGAAGAGGTGCTGAGTGTGACACCTAATGATG 1462
Qy 1474 GCTTTGCTAAAGTCCATTTATGGCTTTCATCTGAAGGCACAGAACAAATTTGCTGAGAGCA 1533
Db 1463 GCTTTGCTAAAGTCCATTTATGGCTTTCATCTGAAGGCACAGAACAAATTTGCTGAGAGCA 1522
Qy 1534 TCCCATATTTAAAGGAAGGATAGAAATCCGAGATCTCGGACCTGATGATGGGAGATTTT 1593
Db 1523 TCCCATATTTAAAGGAAGGATAGAAATCCGAGATCTCGGACCTGATGATGGGAGATTTT 1582
Qy 1594 ATTTCCACCTGGGGATGCCATGAGAGAGGTGGGAAACAAAGAGGCAATATAAGTGGTATG 1653
Db 1583 ATTTCCACCTGGGGATGCCATGAGAGAGGTGGGAAACAAAGAGGCAATATAAGTGGTATG 1642
Qy 1654 AGCTTGGGCAACAGAGAGGACACTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1713
Db 1643 AGCTTGGGCAACAGAGAGGACACTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1702
Qy 1714 ATGGACTGAAAGCACACCTTTGGTGGACCCCAAAAGAAACCGGGCTACACAGAGTTAGTAA 1773
```

Db 1703 ATGGACTGAAGCAGAGCTTTGGTGGACCCCAAAAGAAACGGGCTACACAGAGTTAGTAA 1762  
Qy 1774 AGTCTTTAGAAAGAACTGGAAAGTTAATCCGAGATGAAGGCCCTTGCAGTGTGATGATAAG 1833  
Db 1763 AGTCTTTAGAAAGAACTGGAAAGTTAATCCGAGATGAAGGCCCTTGCAGTGTGATGATAAG 1822  
Qy 1834 CCAAGGCTCTCTCTGCTGAGATGAAGAACTGAGGAAAGAAAGGGAATGAGGAGCAGT 1893  
Db 1823 CCAAGGCTCTCTCTGCTGAGATGAAGAACTGAGGAAAGAAAGGGAATGAGGAGCAGT 1882  
Qy 1894 TCACGCTGGCAGCAGAGGAAAGAAATGAAATGCTGCAAGAGGAGCTCTTAAACCT 1953  
Db 1883 TCACGCTGGCAGCAGAGGAAAGAAATGAAATGCTGCAAGAGGAGCTCTTAAACCT 1942  
Qy 1954 GTACCTTACTAGAAAGTTCCCGAGACACACAGGATGAGAGGAGGACAGATCAAAATATT 2013  
Db 1943 GTACCTTACTAGAAAGTTCCCGAGACACACAGGATGAGAGGAGGACAGATCAAAATATT 2002  
Qy 2014 CCATCATGACCCCGGACTCAGTGTGGCCGACACACAGGCGCCACAACTGCAGGCTCC 2073  
Db 2003 CCATCATGACCCCGGACTCAGTGTGGCCGACACACAGGCGCCACAACTGCAGGCTCC 2062  
Qy 2074 GAATCAGCTGGGCTTGGTGTATCCCAAGAGGCTGCAAGATTCGATGTCACAGAGA 2133  
Db 2063 GAATCAGCTGGGCTTGGTGTATCCCAAGAGGCTGCAAGATTCGATGTCACAGAGA 2122  
Qy 2134 CCA 2136  
Db 2123 CCA 2125

RESULT 2  
BM806632 1050 bp mRNA linear EST 05-MAR-2002  
LOCUS AGENCOURT\_6543030 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5548943  
DEFINITION 5' mRNA sequence.  
ACCESSION BM806632.1 GI:19123455  
VERSION BM806632.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1050)  
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbe-r@mail.nih.gov](mailto:cgapbe-r@mail.nih.gov)  
Tissue Procurement: ATCC/DCTD/DTDP  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM12258 row: k column: 24  
High quality sequence stop: 649.  
Location/Qualifiers  
1. 1050  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5548943"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_72"  
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;  
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 2 kb. Library constructed by Life  
Technologies."

Query Match 33.0%; Score 766; DB 4; Length 1050;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 766; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 865 ATATCTCTGTAGAGATTCACAGGTAATTTGTAGAAAGAAAGTAAAGCATTTTCTCTGTGGAAG 924  
Db 1 ATATCTCTGTAGAGATTCACAGGTAATTTGTAGAAAGAAAGTAAAGCATTTTCTCTGTGGAAG 924  
Qy 925 AACAGCAGGAGTACACACAGAAACAAATAGAAAAACAGATGATCCAGAAACAAAAAGCAA 984  
Db 61 AACAGCAGGAGTACACACAGAAACAAATAGAAAAACAGATGATCCAGAAACAAAAAGCAA 120.  
Qy 985 AAGTTAAGAAAAAGAGCCCTTAACCTTTTAAATTAATTTTGAATAGACTATTTAAAGCTGAAC 1044  
Db 121 AAGTTAAGAAAAAGAGCCCTTAACCTTTTAAATTAATTTTGAATAGACTATTTAAAGCTGAAC 180  
Qy 1045 TTGATGCTGCAGAAAAAAGCTCCGTAAAGGGGAAAAATTTGAGAGAAAGCAATGCAATTTA 1104  
Db 181 TTGATGCTGCAGAAAAAAGCTCCGTAAAGGGGAAAAATTTGAGAGAAAGCAATGCAATTTA 240  
Qy 1105 AAGAACTAGTAGCAAAATACCTTCAGAGTCCACGAGCAAGATATGGGAAGGGCGCAGTGTG 1164  
Db 241 AAGAACTAGTAGCAAAATACCTTCAGAGTCCACGAGCAAGATATGGGAAGGGCGCAGTGTG 300  
Qy 1165 AGGATGATTTGGCTGAGAAAGAGAGAGTAAATGAGGTGCTACGTGGAGCCATCGAGACCT 1224  
Db 301 AGGATGATTTGGCTGAGAAAGAGAGAGTAAATGAGGTGCTACGTGGAGCCATCGAGACCT 360  
Qy 1225 ACCAAGAGTGGCAGGCTACCTGATGCTCCCTGCAGACCTGCTGAAGCTGAGTTGGAAGC 1284  
Db 361 ACCAAGAGTGGCAGGCTACCTGATGCTCCCTGCAGACCTGCTGAAGCTGAGTTGGAAGC 420  
Qy 1285 GTGCTCAGACAGGCAACAAATTTCTAGGTCAATATGAGAGGTTCCTGCTTACCCTGCAGA 1344  
Db 421 GTGCTCAGACAGGCAACAAATTTCTAGGTCAATATGAGAGGTTCCTGCTTACCCTGCAGA 480  
Qy 1345 GATTAGTTCAACTATTTCCTCAATGATCTTCCTTTAAAAAATGACCTTGGCGTGGGATACC 1404  
Db 481 GATTAGTTCAACTATTTCCTCAATGATCTTCCTTTAAAAAATGACCTTGGCGTGGGATACC 540  
Qy 1405 TCTTGATAGAGATATGACATGCAAGAAAGTTTATGAGAGGTGCTGAGTGTGACAC 1464  
Db 541 TCTTGATAGAGATATGACATGCAAGAAAGTTTATGAGAGGTGCTGAGTGTGACAC 600  
Qy 1465 CTATGATGGCTTTGCTAAAGTCCATTTATGGCTTCATCTCTGAAAGGCACAGAAACAAATTTG 1524  
Db 601 CTATGATGGCTTTGCTAAAGTCCATTTATGGCTTCATCTCTGAAAGGCACAGAAACAAATTTG 660  
Qy 1525 CTGAGAGCATCCCATATTATTAAGGAAGGAATAGAAATCCGGAGATCTCTGGCACTGATGATG 1584  
Db 661 CTGAGAGCATCCCATATTATTAAGGAAGGAATAGAAATCCGGAGATCTCTGGCACTGATGATG 720  
Qy 1585 GGAGATTTTATTTCCACCTGGGGGATGCAATGAGAGGGTTGGGAA 1630  
Db 721 GGAGATTTTATTTCCACCTGGGGGATGCAATGAGAGGGTTGGGAA 766

RESULT 3  
BX459083 1085 bp mRNA linear EST 06-MAY-2004  
LOCUS BX459083  
DEFINITION BX459083 Homo sapiens PLACENTA Homo sapiens cDNA clone CS05012YB13  
5-PRIME, mRNA sequence.  
ACCESSION BX459083  
VERSION BX459083.2 GI:47064816  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1085)  
AUTHORS Li, W.B.; Gruber, C.; Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization

QY	762	GATACAGATGATGTAAACATACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAA	821
Db	650	GATACAGATGATGTAAACATACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAA	709
QY	822	AATGAAGGATAGAAATACACAGAACTAATCTCCCTGAGGATAATCTCTGTAGAAAGAT	881
Db	710	AATGAAGGATAGAAATACACAGAACTAATCTCCCTGAGGATAATCTCTGTAGAAAGAT	769
QY	882	TCACAGGTAATGTAGAAAGTAAAGCATTTTCTGTGGAAGAAACAGCAGGAAGTACCA	941
Db	770	TCACAGGTAATGTAGAAAGTAAAGCATTTTCTGTGGAAGAAACAGCAGGAAGTACCA	829
QY	942	CCAGA	946
Db	830	CCAGA	834
RESULT 4			
BM790814			
LOCUS			
DEFINITION			
K-EST0070696 S21SNUS20 Homo sapiens cDNA clone S21SNUS20-16-C08 5',			
mRNA sequence.			
BM790814			
BM790814.1 GI:19139046			
ACCESSION			
EST.			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens			
Homo sapiens (human)			
REFERENCE			
1. (bases 1 to 689)			
Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,			
Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and			
Kim, Y.S.			
TITLE			
21C Frontier Korean EST Project 2001			
JOURNAL			
COMMENT			
Unpublished (2002)			
Contact: Kim YS			
Genome Research Center			
Korea Research Institute of Bioscience & Biotechnology			
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea			
Tel: +82-42-860-4470			
Fax: +82-42-860-4409			
Email: yongsung@mail.kribb.re.kr			
Plate: 16 row: C column: 08			
High quality sequence stop: 689.			
FEATURES			
Location/Qualifiers			
1..689			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="S21SNUS20-16-C08"			
/sex="F"			
/tissue_type="Stomach"			
/cell_type="Floating aggregates"			
/lab_host="Top10F"			
/clone_lib="S21SNUS20"			
/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;			
Site 2: NotI; The poly (A) + RNA was dephosphorylated with			
bacterial alkaline phosphatase (BAP) and then decapped			
with tobacco acid pyrophosphatase (TAP). The decapped			
intact mRNA was ligated with DNA-RNA linker including EcoR			
I site by treatment of T4 RNA ligase and the first strand			
cDNA was synthesized from oligo dt-selected mRNA by			
priming with dt-tailed vector. The dt-tailed vector was			
adjusted to have about 60nt. The cDNA vector was			
circularized with E. coli DNA ligase after digestion of			
EcoRI which site is also included in vector. An RNA strand			
converted to a DNA strand by Okayama-Berg method. The			
obtained cDNA vectors were used for transformation of			
competent cells E. coli Top10F by electroporation method.			
The cDNA libraries constructed by this method are			
full-length enriched cDNA library."			
ORIGIN			

Unpublished (2001)	
On May 22, 2003 this sequence version replaced gi:31027087.	
Contact: Genoscope	
Genoscope - Centre National de Sequencage	
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE	
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr	
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime	
end enriched, double-strand cDNA was digested with Not I and cloned	
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library	
was not normalized. Library was constructed by Life Technologies, a	
division of Invitrogen.	
This sequence belongs to sequence cluster 4537.r	
For more information about this cluster, see	
http://www.genoscope.cns.fr/cdna?s=CS0DE012CA07QP1&c=4537.r.	
Location/Qualifiers	
1..1085	
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/clone="CS0DE012YB13"	
/tissue_type="PLACENTA"	
/clone_lib="Homo sapiens PLACENTA"	
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed	
with a NotI-oligo(dT) primer. Five prime end enriched,	
double-strand cDNA was digested with Not I and cloned into	
the Not I and EcoRV sites of the pCMVSPORT 6 vector.	
Library was not normalized."	
Query Match 31.6%; Score 734; DB 5; Length 1085;	
Best Local Similarity 99.9%; Pred. No. 0;	
Matches 784; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
162 GGAACCTTCATTCCACGGTTTATGGTGATTCGTCGGCGTCTGACATCTGTA	221
50 GGAACCTTCATTCCACGGTTTATGGTGATTCGTCGGCGTCTGACATCTGTA	109
222 GCTGTGCTTGGTTGATCTTGTTCACATATGAGGAAGTTCTAGGAAACTAGGAATCTAT	281
110 GCTGTGCTTGGTTGATCTTGTTCACATATGAGGAAGTTCTAGGAAACTAGGAATCTAT	169
282 GATGCTGATGTCGATGGAGATTTTATGTTGGATGATGCCAAAGTTTATTAGGACTTAAA	341
170 GATGCTGATGTCGATGGAGATTTTATGTTGGATGATGCCAAAGTTTATTAGGACTTAAA	229
342 GAGAGATCTACTTCAGAGCCAGCATGTCGCCAGAGAGGCTGAGCCACACACTGAGCCC	401
230 GAGAGATCTACTTCAGAGCCAGCATGTCGCCAGAGAGGCTGAGCCACACACTGAGCCC	289
402 GAGGACAGGTTCTCTGTGAGGAGAGAGCCAGAGATATCGAAGATGAAGCAAAAGACAA	461
290 GAGGACAGGTTCTCTGTGAGGAGAGAGCCAGAGATATCGAAGATGAAGCAAAAGACAA	349
462 ATTTCAGTCCCTTCTCCATGAATGTTACACGAGACATGTTTGAGGAGAGACTTGCA	521
350 ATTTCAGTCCCTTCTCCATGAATGTTACACGAGACATGTTTGAGGAGAGACTTGCA	409
522 CAAGAAGATGACCCACAGGAGAACCAACAAAGAGGATGATGATTTCTTATGGGACT	581
410 CAAGAAGATGACCCACAGGAGAACCAACAAAGAGGATGATGATTTCTTATGGGACT	469
582 GATGTAGATGATAGATTGAGACCTTGGAACTTGAAGTATCTCATGAAGAAACCGAGCAT	641
470 GATGTAGATGATAGATTGAGACCTTGGAACTTGAAGTATCTCATGAAGAAACCGAGCAT	529
642 AGTTACACCTGGAAGAGACAGTTTTCACAGACTGTAATCAGGATATGAAGAGATGATG	701
530 AGTTACACCTGGAAGAGACAGTTTTCACAGACTGTAATCAGGATATGAAGAGATGATG	589
702 TCTGAGCAGGAAATCCAGATTCAGTGACACAGTATGAGAGATGAAGATTCACACCAT	761
590 TCTGAGCAGGAAATCCAGATTCAGTGACACAGTATGAGAGATGAAGATTCACACCAT	649









LOCUS BP336604 603 bp mRNA linear EST 17-SEP-2004  
 DEFINITION BP336604 Sugano cDNA library, coronary artery smooth muscle cell  
 Homo sapiens cDNA clone SMR06083, mRNA sequence.  
 ACCESSION BP336604  
 VERSION BP336604.1 GI:52266184  
 KEYWORDS EST.  
 ORGANISM Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 603)  
 Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,  
 Mizushima-Sugano, J., Nakai, K. and Sugano, S.  
 Sequence comparison of human and mouse genes reveals a homologous  
 block structure in the promoter regions  
 Genome Res. 14 (9), 1711-1718 (2004)  
 Contact: Yutaka Suzuki  
 Department of Virology  
 Institute of Medical Science, University of Tokyo  
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
 Email: ysuzuki@ims.u-tokyo.ac.jp.  
 Location/Qualifiers  
 1..603  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="SMR06083"  
 /tissue\_type="coronary artery"  
 /cell\_type="smooth muscle cell"  
 /clone\_lib="Sugano cDNA library, coronary artery smooth  
 muscle-cell"  
 FEATURES  
 source  
 1..603  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="SMR06083"  
 /tissue\_type="coronary artery"  
 /cell\_type="smooth muscle cell"  
 /clone\_lib="Sugano cDNA library, coronary artery smooth  
 muscle-cell"  
 ORIGIN  
 Query Match 24.4%; Score 568; DB 5; Length 603;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-293;  
 Matches 568; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 801 GCAGTATGACCTCTGAGAAATGAGGATGAGAAATCAGAGAGTAACTGCTCCCTCCT 860  
 DB 36 GCAGTATGACCTCTGAGAAATGAGGATGAGAAATCAGAGAGTAACTGCTCCCTCCT 95  
 QY 861 GAGGATAATCCTGTAGAGATTACAGGTAATTTAGAGAAAGTAAAGATTTTCTCTGTG 920  
 DB 96 GAGGATAATCCTGTAGAGATTACAGGTAATTTAGAGAAAGTAAAGATTTTCTCTGTG 155  
 QY 921 GAAGAACAGCAGGAAGTACACAGAAACAAATAGAAACAGAGATGATCCAGAACAAAAA 980  
 DB 156 GAAGAACAGCAGGAAGTACACAGAAACAAATAGAAACAGAGATGATCCAGAACAAAAA 215  
 QY 981 GCAAAAGTTAAGAAAAAGAGCCTAACTTTTAAATTAATTTTATAAGACTATTAAAGCT 1040  
 DB 216 GCAAAAGTTAAGAAAAAGAGCCTAACTTTTAAATTAATTTTATAAGACTATTAAAGCT 275  
 QY 1041 GAACCTTGATGCTGCAGAAAACTCCGTAAAGGGGAAAAATTTGAGGAAGCAGTGAATGCA 1100  
 DB 276 GAACCTTGATGCTGCAGAAAACTCCGTAAAGGGGAAAAATTTGAGGAAGCAGTGAATGCA 335  
 QY 1101 TTTAAGAACTAGTACGCAATACCTCAGAGTCCAGAGAGATATGGGAGGGCGCAG 1160  
 DB 336 TTTAAGAACTAGTACGCAATACCTCAGAGTCCAGAGAGATATGGGAGGGCGCAG 395  
 QY 1161 TGTGAGGATGATTTGGCTGAGAGAGAGAGAAATTAAGGTGCTACGTGAGGCATCGAG 1220  
 DB 396 TGTGAGGATGATTTGGCTGAGAGAGAGAGAAATTAAGGTGCTACGTGAGGCATCGAG 455  
 QY 1221 ACCTCAAGAGGTGGCAGCCTACCTGATGTCCTCGAGACCTGCTGAGGCTGAGTTTG 1280  
 DB 456 ACCTCAAGAGGTGGCAGCCTACCTGATGTCCTCGAGACCTGCTGAGGCTGAGTTTG 515  
 QY 1281 AAGCGTCGCTCAGACAGCAACAAATTTCTAGGTCATATGAGAGGTTCCCTGCTTACCCTG 1340  
 DB 516 AAGCGTCGCTCAGACAGCAACAAATTTCTAGGTCATATGAGAGGTTCCCTGCTTACCCTG 575

QY 1341 CAGAGATTAGTTCAACTATTTCCCAATG 1368  
 DB 576 CAGAGATTAGTTCAACTATTTCCCAATG 603  
 RESULT 10  
 BP364957  
 LOCUS BP364957  
 DEFINITION BP364957 Sugano cDNA library, fetal lung fibroblast TIG Homo  
 sapiens cDNA clone TIR05529, mRNA sequence.  
 ACCESSION BP364957  
 VERSION BP364957.1 GI:52295202  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 583)  
 Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,  
 Mizushima-Sugano, J., Nakai, K. and Sugano, S.  
 Sequence comparison of human and mouse genes reveals a homologous  
 block structure in the promoter regions  
 Genome Res. 14 (9), 1711-1718 (2004)  
 Contact: Yutaka Suzuki  
 Department of Virology  
 Institute of Medical Science, University of Tokyo  
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan  
 Email: ysuzuki@ims.u-tokyo.ac.jp.  
 Location/Qualifiers  
 1..583  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="TIR05529"  
 /tissue\_type="lung"  
 /cell\_type="fibroblast"  
 /cell\_line="TIG"  
 /dev\_stage="fetal"  
 /clone\_lib="Sugano cDNA library, fetal lung fibroblast  
 TIG"  
 FEATURES  
 source  
 1..583  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="TIR05529"  
 /tissue\_type="lung"  
 /cell\_type="fibroblast"  
 /cell\_line="TIG"  
 /dev\_stage="fetal"  
 /clone\_lib="Sugano cDNA library, fetal lung fibroblast  
 TIG"  
 ORIGIN  
 Query Match 24.2%; Score 563; DB 5; Length 583;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-290;  
 Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1102 TTAAGAAGTACGTACGCAATACCTCAGAGTCCAGCAAGATATGGAGCGGCAGT 1161  
 DB 21 TTAAGAAGTACGTACGCAATACCTCAGAGTCCAGCAAGATATGGAGCGGCAGT 80  
 QY 1162 GTGAGGATGATTTGGCTGAGAGAGAGAGAAATTAAGGTGCTACGTGGAGCCATCGAG 1221  
 DB 81 GTGAGGATGATTTGGCTGAGAGAGAGAGAAATTAAGGTGCTACGTGGAGCCATCGAG 140  
 QY 1222 CCTACCAAGAGTGGCCAGCCTACCTGATGTCCTCGAGACCTGCTGAAGCTGAGTTGA 1281  
 DB 141 CCTACCAAGAGTGGCCAGCCTACCTGATGTCCTCGAGACCTGCTGAAGCTGAGTTGA 200  
 QY 1282 AGCGTCGCTCAGAGCGCAACAATTTCTAGGTCATATGAGAGGTTCCCTGCTTACCCTGC 1341  
 DB 201 AGCGTCGCTCAGAGCGCAACAATTTCTAGGTCATATGAGAGGTTCCCTGCTTACCCTGC 260  
 QY 1342 AGAGATTAGTTCACCTATTTCCCAATGATCTTCTTAAAAAATGACCTTGGCGTGGAT 1401  
 DB 261 AGAGATTAGTTCACCTATTTCCCAATGATCTTCTTAAAAAATGACCTTGGCGTGGAT 320  
 QY 1402 ACCTCTTGATAGGAGATAATGACATGCAATGCAAGAAAGTTTATGAGAGGCTGAGTGA 1461  
 DB 321 ACCTCTTGATAGGAGATAATGACATGCAATGCAAGAAAGTTTATGAGAGGCTGAGTGA 380  
 QY 1462 CACCTAATGATGGCTTTGCTAAAGTCCATTTATGGCTTCATCTGAAGGCACAGAACAAAA 1521  
 DB 381 CACCTAATGATGGCTTTGCTAAAGTCCATTTATGGCTTCATCTGAAGGCACAGAACAAAA 440

```
QY 1522 TTCTGAGAGCATCCCATATTTAAAGGAAGAAATAGATCCGAGATCCTGGCACTGATG 1581
Db 441 TTCTGAGAGCATCCCATATTTAAAGGAAGAAATAGATCCGAGATCCTGGCACTGATG 500
QY 1582 ATGGGAGATTTTATTTCCACCTGGGGATGCCATGACAGAGGCTTGGGAACAAGAGGCAT 1641
Db 501 ATGGGAGATTTTATTTCCACCTGGGGATGCCATGACAGAGGCTTGGGAACAAGAGGCAT 560
QY 1642 ATAAGTGGTATGAGCTTGGGCAC 1664
Db 561 ATAAGTGGTATGAGCTTGGGCAC 583

RESULT 11
AUI10952
DEFINITION AUI10952 NT2RP3 Homo sapiens cDNA clone NT2RP3001731 5', mRNA
sequence.
ACCESSION AUI10952
VERSION AUI10952
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Ota, I., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,
Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and
Isogai, T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- and 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
source
1. 859
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RP3001731"
/cell_type="tetatocarcinoma"
/cell_line="NT2"
/clone_lib="NT2RP3"
/notes="Vector: pME18SFL3; mRNA from NT2 neuronal precursor
cells after 2-weeks retinoic acid (RA) induction"

Query Match 23.9%; Score 556; DB 1; Length 859;
Best Local Similarity 99.7%; Pred. No. 8.8e-287;
Matches 656; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1013 AAATPAAATTTGATAAGACTTAAAGCTGAACCTGATGCTGCAGAAAAAATCCGTAAGAAG 1072
Db 1 AAATPAAATTTGATAAGACTTAAAGCTGAACCTGATGCTGCAGAAAAAATCCGTAAGAAG 60
QY 1073 GGGAAAAATTTGAGGAGCAGTGAATGATTTAAAGACTAGTACGCAATACCTTCAGAG 1132
Db 61 GGGAAAGTTTGAGGAGCAGTGAATGATTTAAAGACTAGTATGCAATACCTTCAGAG 120
QY 1133 TCCACGAGCAAGATATGGAAGCCGACGTGTGAGGATGATTTGGCTGAGAGAGAGAG 1192
Db 121 TCCACGAGCAAGATATGGAAGCCGACGTGTGAGGATGATTTGGCTGAGAGAGAGAG 180
QY 1193 TAATGAGGTGCTACGTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCTACCTGATGT 1252
```

```
Db 181 TAATGAGGTGCTACGTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCTACCTGATGT 240
QY 1253 CCTCTCAGACCTGCTGAAGCTGAGTTTGAAGCGTCGCTCAGACAGGCAACAATTTCTAGG 1312
Db 241 CCTCTCAGACCTGCTGAAGCTGAGTTTGAAGCGTCGCTCAGACAGGCAACAATTTCTAGG 300
QY 1313 TCATATGAGAGGTTTCCCTGCTTACCTCGCAGAGATTAGTTCAAACTATTTCCTCAATGATAC 1372
Db 301 TCATATGAGAGGTTTCCCTGCTTACCTCGCAGAGATTAGTTCAAACTATTTCCTCAATGATAC 360
QY 1373 TTCCTTTAAAAAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATATGACATCAAA 1432
Db 361 TTCCTTTAAAAAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATATGACATCAAA 420
QY 1433 GAAAGTTTATGAGAGGTCGCTGAGTGTGACACTAATGATGCTTCTTAAAGTCCATTA 1492
Db 421 GAAAGTTTATGAGAGGTCGCTGAGTGTGACACTAATGATGCTTCTTAAAGTCCATTA 480
QY 1493 TGGCTTATCTCTGAGGCGACAGACAAAATTGCTGAGAGCATCCCATATTTAAAGGAAGG 1552
Db 481 TGGCTTATCTCTGAGGCGACAGACAAAATTGCTGAGAGCATCCCATATTTAAAGGAAGG 540
QY 1553 AATAGAATCCGGAGATCCCTGGCACTGATGGGAGATTTTATTTCCACCTGGGGATGC 1612
Db 541 AATAGAATCCGGAGATCCCTGGCACTGATGGGAGATTTTATTTCCACCTGGGGATGC 600
QY 1613 CATGCAGAGGTTTGGGAACAAGAGGCATATAAGTGTGATGAGCTTGGGCAAGAGA 1670
Db 601 CATGCAGAGGTTTGGGAACAAGAGGCATATAAGTGTGATGAGCTTGGGCAAGAGA 658

RESULT 12
AW996828/c
LOCUS AW996828
DEFINITION QV3-BN0047-160300-123-a02 BN0047 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW996828
VERSION AW996828.1 GI:8257062
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Dias Neto, E., Garcia Correa, R., Verjovetki-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?i=6t2-QV3-BN0047-160
300-123-a02&t3=2000-03-16&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 643.
FEATURES
source
1. 651
/organism="Homo sapiens"
/mol_type="mRNA"
```

On May 13, 2003 this sequence version replaced gi:30635043.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
was not normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
This sequence belongs to sequence cluster 4537.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?c=CL0BA007ZC04RPI&c=4537.r.  
Location/Qualifiers  
1. 1011  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CL0BA007ZC04"  
/tissue="PLACENTA"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo (dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

ORIGIN		Query Match	23.4%; Score 543; DB 2; Length 651;
		Best Local Similarity	100.0%; Pred. No. 8.3e-280;
		Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1366	ATGATCTCTTCTTAAATAATGACCTTGGCGTGGATACCTCTTGATAGGAGATAATGACA	1425
DB	620	ATGATCTCTTCTTAAATAATGACCTTGGCGTGGATACCTCTTGATAGGAGATAATGACA	561
QY	1426	ATGCAAGAAAGTTTATGAAGAGTGTGAGTGTGACACCTAATGATGGCTTTGCTAAAG	1485
DB	560	ATGCAAGAAAGTTTATGAAGAGTGTGAGTGTGACACCTAATGATGGCTTTGCTAAAG	501
QY	1486	TCCATTATGCTTCTATCTGAGGACACAGACAAATTTGCTGAGAGCATCCCATATTAA	1545
DB	500	TCCATTATGCTTCTATCTGAGGACACAGACAAATTTGCTGAGAGCATCCCATATTAA	441
QY	1546	AGGAAGGAATAGAAATCCGAGATCCTGGCACTGATGGAGATTTTATTTCCACTGG	1605
DB	440	AGGAAGGAATAGAAATCCGAGATCCTGGCACTGATGGAGATTTTATTTCCACTGG	381
QY	1606	GGGATGCCATGACAGAGGTGGGAACAAAGAGGCATATAAGTGTATGAGCTTGGGCACA	1665
DB	380	GGGATGCCATGACAGAGGTGGGAACAAAGAGGCATATAAGTGTATGAGCTTGGGCACA	321
QY	1666	AGAGAGGACACTTTCATCTGTCTGCAACGCTCACTCTACAAATGTGAATGGACTGAAAG	1725
DB	320	AGAGAGGACACTTTCATCTGTCTGCAACGCTCACTCTACAAATGTGAATGGACTGAAAG	261
QY	1726	CACAGCTTTGGTGGACCCCAAGAAACGGGTACACAGATTAAGTCTTTAGNAA	1785
DB	260	CACAGCTTTGGTGGACCCCAAGAAACGGGTACACAGATTAAGTCTTTAGNAA	201
QY	1786	GAACTGGAAGTTAATCCAGATGAGGCTTGCAGTGTGATGATAAGCCAAAGTCTCT	1845
DB	200	GAACTGGAAGTTAATCCAGATGAGGCTTGCAGTGTGATGATAAGCCAAAGTCTCT	141
QY	1846	TCTGCTGAGGATGAAACCTGAGGAAAGGGGACTGGAGCCAGTTACGCTGTGGC	1905
DB	140	TCTGCTGAGGATGAAACCTGAGGAAAGGGGACTGGAGCCAGTTACGCTGTGGC	81
QY	1906	AGC 1908	
DB	80	AGC 78	
RESULT 13			
BX403636			
LOCUS			
DEFINITION			
BX403636 Homo sapiens PLACENTA Homo sapiens cDNA clone CL0BA007ZC04			
5-PRIME, mRNA sequence.			
ACCESSION			
BX403636			
VERSION			
BX403636.2 GI:46924400			
KEYWORDS			
EST.			
SOURCE			
Homo sapiens (human)			
ORGANISM			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
1 (bases 1 to 1011)			
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.			
Full-length cDNA libraries and normalization			
Unpublished (2001)			

COMMENT

On May 13, 2003 this sequence version replaced gi:30635043.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
was not normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
This sequence belongs to sequence cluster 4537.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?c=CL0BA007ZC04RPI&c=4537.r.  
Location/Qualifiers  
1. 1011  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CL0BA007ZC04"  
/tissue="PLACENTA"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo (dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

ORIGIN

Query Match	23.3%;	Score 541;	DB 5;	Length 1011;
Best Local Similarity	99.3%;	Pred. No. 1.1e-278;		
Matches 841;	Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0;
QY	1	CGGACCGTGCATGGCCAGCGTAAGATGCCAAGAGCAGCGCCACAGCAGCAGCAGCG	60	
DB	110	CGGACCGTGCATGGCCAGCGTAAGATGCCAAGAGCAGCGCCACAGCAGCAGCAGCG	169	
QY	61	GCTCCGCGCAGCGTAGCACAGTGGCGGCGAGCAGCCCCGGGGCCCGGAGAGAGACAA	120	
DB	170	GCTCCGCGCAGCGTAGCACAGTGGCGGCGAGCAGCCCCGGGGCCCGGAGAGAGACAA	229	
QY	121	AGCATGGAGACACAAAGATGGGAGAAAGCGGAGCTCTCGGAACTTCACTTTCACGT	180	
DB	230	AGCATGGAGACACAAAGATGGGAGAAAGCGGAGCTCTCGGAACTTCACTTTCACGT	289	
QY	181	GGTTTATGGTGTGATTCGATCTGCTGGCGCTCTGGACATCTGTAGCTGTCTGTTGGTTGATC	240	
DB	290	GGTTTATGGTGTGATTCGATCTGCTGGCGCTCTGGACATCTGTAGCTGTCTGTTGGTTGATC	349	
QY	241	TTGTTGACTATGAGGAAGTTCTAGGAAACTTAGGAATCTATGATGCTGATGGATGGAG	300	
DB	350	TTGTTGACTATGAGGAAGTTCTAGGAAACTTAGGAATCTATGATGCTGATGGATGGAG	409	
QY	301	ATTTTGTATGGATGATGCCAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTCAGAGC	360	
DB	410	ATTTTGTATGGATGATGCCAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTCAGAGC	469	
QY	361	CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGCAGGCTTCTCTGTGG	420	
DB	470	CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGCAGGCTTCTCTGTGG	529	
QY	421	AGGCAGAACCCAGCAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCCATG	480	
DB	530	AGGCAGAACCCAGCAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCCATG	589	
QY	481	AAATGGTACACGCGAACAATGTTGAGGAGAAAGACTTGCACACAGAGATGGACCCACAG	540	
DB	590	AAATGGTACACGCGAACAATGTTGAGGAGAAAGACTTGCACACAGAGATGGACCCACAG	649	
QY	541	GAGACACCAACAAGAGGATGATGAGTTCTTATGGCGACTGATGTAGATGATGATTTG	600	
DB	650	GAGAACCAACAAGAGGATGATGAGTTCTTATGGCGACTGATGTAGATGATGATTTG	709	
QY	601	AGACCCCTGGAACCTGAGTATCTCATGAAGAAACCGGACATAGTTACCACCTGGAAGAGA	660	

```

Db      710 AGACCTTGAACCTGAAGTATCTATGAGAAACCGAGCATAGTTTACCACCTGGAAGAGA 769
Qy      661 CAGTTTCAACAAGCTGTAAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720
Db      770 CAGTTTCAACAAGCTGTATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 829
Qy      721 ATTCCAGTGAACCACTAGTGAAGATGAAGATGGAAGATGGAAGATGGAAGATGGAAGAT 780
Db      830 ATTCCAGTGAACCACTAGTGAAGATGAAGATGGAAGATGGAAGATGGAAGATGGAAGAT 889
Qy      781 ACCAAGTCTATGAGAACCAAGCTAGTATATGAACCTCTAGAAATGAAGGATGAATCA 840
Db      890 ACCAAGTCTATGAGAACCAAGCTAGTATATGAACCTCTAGAAATGAAGGATGAATCA 949
Qy      841 CAGAAAT 847
Db      950 CAGAAAT 956

```

```

RESULT 14
CN258117 589 bp mRNA linear EST 16-MAY-2004
LOCUS 1700531862302 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
DEFINITION CN258117
ACCESSION CN258117.1 GI:47274531
VERSION EST.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 589)
Li, Y., Xu, C., Fang, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Lebkowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 589 Std Error: 0.00.
Location/Qualifiers
1. 589
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN_EB"
/notes="oligo dt primed, full-length enriched cDNA library
from embryoid body outgrowths derived from hES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."

```

## FEATURES

```

source
1. 589
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN_EB"
/notes="oligo dt primed, full-length enriched cDNA library
from embryoid body outgrowths derived from hES cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."

```

## ORIGIN

```

Query Match 23.1%; Score 538; DB 7; Length 589;
Best Local Similarity 100.0%; Pred. No. 4e-277;
Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1216 TCGAGACCTACCAAGAGTGGCCAGCTACCTGATGTCCTGAGACCTGCTGAAGCTGA 1275
Db 52 TCGAGACCTACCAAGAGTGGCCAGCTACCTGATGTCCTGAGACCTGCTGAAGCTGA 111
Qy 1276 GTTTGAAGCTTCCTCAGACGCAACAAATTTCTAGTTCATATGAGAGTTCCCTGCTTA 1335
Db 112 GTTTGAAGCTTCCTCAGACGCAACAAATTTCTAGTTCATATGAGAGTTCCCTGCTTA 171
Qy 1336 CCCTGCAGAGATTAGTTCAACTATTTCCTCAATGATATCTTCCATAAATAAGTACCTTGGCG 1395

```

```

Db      172 CCTCAGAGATAGTTCAACTATTTCCTCAATGATGATCTTCTTAAATAAGTCTGGCG 231
Qy      1396 TGGGATACCTCTTATAGGAGATATGACAATGCAAAAGAAAGTTTATGAAGAGTGTCTGA 1455
Db      232 TGGGATACCTCTTATAGGAGATATGACAATGCAAAAGAAAGTTTATGAAGAGTGTCTGA 291
Qy      1456 GTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTATGCTTCATCTCTGAAGGCACAGA 1515
Db      292 GTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTATGCTTCATCTCTGAAGGCACAGA 351
Qy      1516 ACAAATATGCTGAGAGCATCCCATATTTAAAGAGAGGATAGATCCGGAGATCTCGGCA 1575
Db      352 ACAAATATGCTGAGAGCATCCCATATTTAAAGAGAGGATAGATCCGGAGATCTCGGCA 411
Qy      1576 CTGATGATGGGAGATTTTATTTCCACCTGGGGATGCCATGACAGAGGTTGGGAAACAAAG 1635
Db      412 CTGATGATGGGAGATTTTATTTCCACCTGGGGATGCCATGACAGAGGTTGGGAAACAAAG 471
Qy      1636 AGGCATATAAGTGTATGAGCTTGGGCAACAAGAGGAGACACTTTGCACTCTGTCTGGCAAC 1695
Db      472 AGGCATATAAGTGTATGAGCTTGGGCAACAAGAGGAGACACTTTGCACTCTGTCTGGCAAC 531
Qy      1696 GCTCACTCTCAATGTGAATGGAATGGAAGCAGACGCTTGTGACCCCAAAAGAAAC 1753
Db      532 GCTCACTCTCAATGTGAATGGAATGGAAGCAGACGCTTGTGACCCCAAAAGAAAC 589

RESULT 15
BP266690 581 bp mRNA linear EST 16-SEP-2004
LOCUS BP266690 Sugano cDNA library, thyroid JTH Homo sapiens cDNA clone
DEFINITION JTH04779, mRNA sequence.
ACCESSION BP266690
VERSION BP266690.1 GI:52181921
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 581)
Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
Location/Qualifiers
1. 581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="JTH04779"
/tissue_type="thyroid"
/cell_line="JTH"
/clone_lib="Sugano cDNA library, thyroid JTH"
/notes="Thyroid tumor"

ORIGIN
Query Match 23.1%; Score 537; DB 5; Length 581;
Best Local Similarity 100.0%; Pred. No. 1.4e-276;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1158 CAGTGTGAGATGATTGGCTGAGAGAGGAGAGTAATGAGTGTCTACGTGGAGCCATC 1217
Db 45 CAGTGTGAGATGATTGGCTGAGAGAGGAGAGTAATGAGTGTCTACGTGGAGCCATC 104
Qy 1218 GAGACCTACCAAGAGTGGCCAGCTACCTGATGTCCTTGCAGACCTCTCTGAAGCTGAGT 1277

```

Db	105	GAGACCTTACCAAGAGGTGGCCAGCTTACCTGATGTCCCTGCAGACCTGCTGAAGCTGAGT	164
Qy	1278	TTGAAGCGTGGCTCAGACAGGCAACAAATTTCTAGGTCAATATGAGAGGTTCCTGCTTACC	1337
Db	165	TTGAAGCGTGGCTCAGACAGGCAACAAATTTCTAGGTCAATATGAGAGGTTCCTGCTTACC	224
Qy	1338	CTGCAGAGATGAGTTCAACTATTTCCCAATGATACCTTTAAATAATGACCTTGGCGTG	1397
Db	225	CTGCAGAGATGAGTTCAACTATTTCCCAATGATACCTTTAAATAATGACCTTGGCGTG	284
Qy	1398	GGATACCTCTTGATAGAGAGTAATGACAATGCCAAGAAAGTTTATGAAGAGGTGCTGAGT	1457
Db	285	GGATACCTCTTGATAGAGAGTAATGACAATGCCAAGAAAGTTTATGAAGAGGTGCTGAGT	344
Qy	1458	GTGCACACTAATGATGGCTTGTCTAAAGTCAATATGCTTATGCTTCCCTGAAGGACAGAAC	1517
Db	345	GTGCACACTAATGATGGCTTGTCTAAAGTCAATATGCTTATGCTTCCCTGAAGGACAGAAC	404
Qy	1518	AAATTTGCTGAGAGCATCCCATATTTAAAGGAAGGAATAGAAATCCGGAGATCCTGGCACT	1577
Db	405	AAATTTGCTGAGAGCATCCCATATTTAAAGGAAGGAATAGAAATCCGGAGATCCTGGCACT	464
Qy	1578	GATGATGGAGATTTTATTTCCACCTGGGGGATGCCATGACAGAGGTTGGGAACAAAGAG	1637
Db	465	GATGATGGAGATTTTATTTCCACCTGGGGGATGCCATGACAGAGGTTGGGAACAAAGAG	524
Qy	1638	GCATATAAGTGGTATGAGCTTGGGCAACAGAGAGGACATTTGTCATCTGTCTGGCAA	1694
Db	525	GCATATAAGTGGTATGAGCTTGGGCAACAGAGAGGACATTTGTCATCTGTCTGGCAA	581
RESULT 16			
LOCUS	BP270196	582 bp	mRNA linear EST 16-SEP-2004
DEFINITION	BP270196 Sugano cDNA library, small intestine Homo sapiens CDNA clone KAR01529, mRNA sequence.		
ACCESSION	BP270196		
VERSION	BP270196.1	GI:52185428	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 582)		
AUTHORS	Suzuki,Y., Yanashita,R., Shiota,M., Sakakibara,Y., Chiba,J., Mizushima-Sugano,J., Nakai,K. and Sugano,S.		
TITLE	Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions		
JOURNAL	Genome Res. 14 (9), 1711-1718 (2004)		
COMMENT	Contact: Yutaka Suzuki Department of Virology Institute of Medical Science, University of Tokyo 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan Email: yusuzuki@ims.u-tokyo.ac.jp.		
FEATURES	Location/Qualifiers		
Source	1..582		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="KAR01529"		
	/tissue_type="small intestine"		
	/clone_lib="Sugano cDNA library, small intestine"		
ORIGIN			
Query Match	22.8%;	Score 531,	DB 5: Length 582;
Best Local Similarity	99.8%;	Pred. No. 2.3e-273;	
Matches 581; Conservative	0;	Mismatches 1;	Indels 0; Gaps 0;
Qy	1651	ATGAGCTTGGCCACAGAGAGGACACTTTTGCATCTGCTGGCAACGCTCACTCAATG	1710
Db	1	ATGAGCTTGGCCACAGAGAGGACACTTTTGCATCTGCTGGCAACGCTCACTCAATG	60
Qy	1711	TGAATGACCTGAAAGACACAGCTTTGGTGGACCCCAAGAAAGAGGGCTACAGAGTTAG	1770
Db	61	TGAATGACCTGAAAGACACAGCTTTGGTGGACCCCAAGAAAGAGGGCTACAGAGTTAG	120
Qy	1771	TAAAGTCTTTAGAAAGAACTGGAAGTTAAATCCGAGATGAAGGCCCTTCAGTGTGATGATA	1830
Db	121	TAAAGTCTTTAGAAAGAACTGGAAGTTAAATCCGAGATGAAGGCCCTTCAGTGTGATGATA	180
Qy	1831	AAGCCAAAGGCTCTTTCCTGCTGAGGATGAATAAATCCTGAGGAAAAAGGGGAGTGGAGCC	1890
Db	181	AAGCCAAAGGCTCTTTCCTGCTGAGGATGAATAAATCCTGAGGAAAAAGGGGAGTGGAGCC	240
Qy	1891	AGTTACGCTGTGGCAGCAGGAAGAAGAAATGAAATGCAATGCTGCAAGAGAGCTCCTAAAA	1950
Db	241	AGTTACGCTGTGGCAGCAGGAAGAAGAAATGAAATGCAATGCTGCAAGAGAGCTCCTAAAA	300
Qy	1951	CCTGTACCTTTACTAGAAAAGTTTCCCGAGACAACAGGATGCAGAAAGAGGACAGATCAAT	2010
Db	301	CCTGTACCTTTACTAGAAAAGTTTCCCGAGACAACAGGATGCAGAAAGAGGACAGATCAAT	360
Qy	2011	ATTCCATCATGCAACCCCGGACTCAGCTGTGGCCGCAACAGGGCCCAAACTGCAGGC	2070
Db	361	ATTCCATCATGCAACCCCGGACTCAGCTGTGGCCGCAACAGGGCCCAAACTGCAGGC	420
Qy	2071	TCCGAATGCACCTGGGCTTGGTATTTCCCAAGAAAGGCTGCAAGATTCGATGTGCCAAG	2130
Db	421	TCCGAATGCACCTGGGCTTGGTATTTCCCAAGAAAGGCTGCAAGATTCGATGTGCCAAG	480
Qy	2131	AGACCAAGACCTGGGAGGAAGCAAGGTGCTCATCTTTGATGATCTCTTTGATGATCTCTTT	2190
Db	481	AGACCAAGACCTGGGAGGAAGCAAGGTGCTCATCTTTGATGATCTCTTTGATGATCTCTTT	2232
Qy	2191	TATGGCAGGATGCTCATCTTTCCGGCTGATTTATTCATCTGGG	2232
Db	541	TATGGCAGGATGCTCATCTTTCCGGCTGATTTATTCATCTGGG	582
RESULT 17			
LOCUS	CD643948	765 bp	mRNA linear EST 17-JUN-2003
DEFINITION	AGENCOURT 1450808 NIA Human H1 Embryonic Stem Cell cDNA library (Long) Homo sapiens cDNA clone IMAGE:30425975 5', mRNA sequence.		
ACCESSION	CD643948		
VERSION	CD643948.1	GI:31815344	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 765)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: cgapbs-re@mail.nih.gov Tissue Procurement: Irene Ginis and Mahendra Rao, NIA cDNA Library Preparation: Yulan Piao and Minoru KO CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC c lone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: NDAM511 row: a column: 24 High quality sequence stop: 605. Location/Qualifiers		
FEATURES	1..765		
Source	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:30425975"		
	/tissue_type="Embryonic Stem cells"		





Db	241	GGATGATTGGCTGAGAAGAGGAGAAAGTAATGAGGTGCTAGCTGGAGCCATCGAGACCTA	300
Qy	1226	CCAAGAGGTGGCCAGCCTACTGATGTCCTCGCAGACCTGCTGAAGCTGAGTTGAAGCG	1285
Db	301	CCAAGAGGTGGCCAGCCTACTGATGTCCTCGCAGACCTGCTGAAGCTGAGTTGAAGCG	360
Qy	1286	TCGCTCAGACGAGCAACAATTTCTAGTCTATGATGAGAGGTTCCCTGCTTACCTCGCAGAG	1345
Db	361	TCGCTCAGACGAGCAACAATTTCTAGTCTATGATGAGAGGTTCCCTGCTTACCTCGCAGAG	420
Qy	1346	ATTAGTTCAACTATTTCCCAATGATCTTCTTAAAAAATGACCTTGGCGCTGGGATACCT	1405
Db	421	ATTAGTTCAACTATTTCCCAATGATCTTCTTAAAAAATGACCTTGGCGCTGGGATACCT	480
Qy	1406	CTTGATAGGAGATATGACAAATGCAAAAGAAAGTTTAT	1442
Db	481	CTTGATAGGAGATATGCAATGCAAAAGAAAGTTTAT	517
RESULT 19			
LOCUS	BI333062	721 bp	mRNA linear EST 30-JUL-2001
DEFINITION	602996326F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5138622 5',		
ACCESSION	BI333062		mRNA sequence.
VERSION	BI333062.1	GI:15017719	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: Incyte Genomics, Inc.		
	Clone sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: LLAW1341 row: c column: 07		
	High quality sequence stop: 719.		
FEATURES	Location/Qualifiers		
source	1..721		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5138622"		
	/tissue type="cervical carcinoma cell line"		
	/lab host="DH10B"		
	/clone lib="NIH MGC 12"		
	/note="Organ: Cervix; Vector: pCMV-SPORT6; Site 1: NotI;		
	Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.		
	Average insert size 1.4 Kb. Library prepared by Life		
	Technologies."		
ORIGIN			
Query Match	21.8%;	Score 507;	DB 4; Length 721;
Best Local Similarity	99.8%;	Pred. No. 2e-260;	
Matches 557;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
Qy	1	CGGACCGTGCATGGCCCGGTAAGATCCCAAGAGAGCGCGCAACACGACGAGCG	60
Db	113	CGGACCGTGCATGGCCCGGTAAGATCCCAAGAGAGCGCGCAACACGACGAGCG	172
Qy	61	GCTCCGCGCAGCGTAGCAGAGTGGCGGAGCAGAGCCCGGGCGCCCGGAGACACAA	120
Db	173	GCTCCGCGCAGCGTAGCAGAGTGGCGGAGCAGAGCCCGGGCGCCCGGAGACACAA	232

Qy	121	AGCATGGAGGACACAAAGAAATGGAGGAAAGCGGACTCTCGGAACTTTCATCTTTCACGT	180
Db	233	AGCATGGAGGACACAAAGAAATGGAGGAAAGCGGACTCTCAGGAACCTTCATCTTTCACGT	292
Qy	181	GGTTTATGGTATGATTCATGTTGGCGCTGCGACATCTGTAGCTGCTGTTGGTTTGATC	240
Db	293	GGTTTATGGTATGATTCATGTTGGCGCTGCGACATCTGTAGCTGCTGTTGGTTTGATC	352
Qy	241	TTGTTGACTATGAGGAAGTTCTAGGAAAATAGGAATCTATGATCTGATGGTATGAGAG	300
Db	353	TTGTTGACTATGAGGAAGTTCTAGGAAAATAGGAATCTATGATCTGATGGTATGAGAG	412
Qy	301	ATTTTATGATGATGATGCTCAAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTCAGAGC	360
Db	413	ATTTTATGATGATGATGCTCAAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTCAGAGC	472
Qy	361	CAGCAGTCCCGCAGAGAGGCTGAGCCACACACACTGAGCCGAGGAGGAGGTTCTGTGG	420
Db	473	CAGCAGTCCCGCAGAGAGGCTGAGCCACACACACTGAGCCGAGGAGGAGGTTCTGTGG	532
Qy	421	AGCGAAGCCCGCAGAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCCATG	480
Db	533	AGCGAAGCCCGCAGAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCCATG	592
Qy	481	AAATGTTACACCGCAGAACATGTTGAGGGAGAGACTTGCACAAAGAGATGAGCCACAG	540
Db	593	AAATGTTACACCGCAGAACATGTTGAGGGAGAGACTTGCACAAAGAGATGAGCCACAG	652
Qy	541	GAGAACCAACACAGAGG 558	
Db	653	GAGAACCAACACAGAGG 670	
RESULT 20			
LOCUS	CN368832	711 bp	mRNA linear EST 16-MAY-2004
DEFINITION	17000418219954 GRN_EB Homo sapiens cDNA 5', mRNA sequence.		
ACCESSION	CN368832		
VERSION	CN368832.1	GI:47368766	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
	Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,		
	Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,		
	Lebkowski, J. and Stanton, L.W.		
	Transcriptome characterization elucidates signaling networks that		
	control human ES cell growth and differentiation		
	Nat. Biotechnol. 22 (6), 707-716 (2004)		
	Contact: Brandenberger R		
	Regenerative Medicine		
	Geron Corporation		
	230 Constitution Drive, Menlo Park, CA 94025, USA		
	Tel: 650 473 8658		
	Fax: 650 473 7760		
	Email: rbrandenberger@geron.com		
	Insert Length: 711 Std Error: 0.00.		
FEATURES	Location/Qualifiers		
source	1..711		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/tissue type="embryonic stem cells, embryoid bodies		
	derived from H1, H7 and H9 cells"		
	/clone lib="GRN_EB"		
	/note="Oligo dt primed, full-length enriched cDNA library		
	from embryoid body outgrowths derived from HES cell lines		
	H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free		
	conditions."		
ORIGIN			

Query Match 21.2%; Score 493; DB 7; Length 711;  
 Best Local Similarity 99.8%; Pred. No. 6.8e-253;  
 Matches 543; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGACCGTCCCAATGCCCGCAGCGTAAAGATCCCAAGAGCGCGCAACAGCAGCAGCAGCG 60  
 DB 168 CGGACCGTCCCAATGCCCGCAGCGTAAAGATCCCAAGAGCGCGCAACAGCAGCAGCAGCG 227  
 QY 61 GCTCCGGCAGCGGTAGCAGCAGTCCGGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 120  
 DB 228 GCTCCGGCAGCGGTAGCAGCAGTCCGGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 287  
 QY 121 AGCATGGAGGACACAGATGGGAGGAAAGCGGAGCTCTCGGAACTTCATCTTCACGT 180  
 DB 288 AGCATGGAGGACACAGATGGGAGGAAAGCGGAGCTCTCGGAACTTCATCTTCACGT 347  
 QY 181 GGTATGCTGATGTCATTTGCTGGGCTCTGGACATCTGTAGCTGTGTTGTTGATC 240  
 DB 348 GGTATGCTGATGTCATTTGCTGGGCTCTGGACATCTGTAGCTGTGTTGATC 407  
 QY 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGCTGATGCGAG 300  
 DB 408 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGCTGATGCGAG 467  
 QY 301 ATTGTTGATGCTGATGTCATTTGCTGGGCTCTGGACATCTGTAGCTGTGTTGATC 360  
 DB 468 ATTGTTGATGCTGATGTCATTTGCTGGGCTCTGGACATCTGTAGCTGTGTTGATC 527  
 QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTGTTGG 420  
 DB 528 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTGTTGG 587  
 QY 421 AGGCAAGACCCAGCAATATCGAAGATGAAGCAAAAGAACTTCACTGCTCTTCCATG 480  
 DB 588 AGGCAAGACCCAGCAATATCGAAGATGAAGCAAAAGAACTTCACTGCTCTTCCATG 547  
 QY 481 AAATGGTACGCGACACATGTTGAGGAGAGACTTCCACAGAGAGATGGGCCACAG 540  
 DB 648 AAATGGTACGCGACACATGTTGAGGAGAGACTTCCACAGAGAGATGGGCCACAG 707  
 QY 541 GAGA 544  
 DB 708 GAGA 711

## RESULT 21

BQ232651  
 LOCUS AGENCOURT 7574911 NIH\_MGC\_68 Homo sapiens cDNA clone IMAGE:6059910  
 5', mRNA sequence.

ACCESSION BQ232651  
 VERSION BQ232651.1 GI:20414051

KEYWORDS EST.  
 SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC http://mgi.nci.nih.gov/.

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cga@bbs-t@mail.nih.gov

Tissue Procurement: DCTP/DTP/Gazdar

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing Arrayed by: The I.M.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM13327 row: f column: 07

High quality sequence stop: 667.

## FEATURES

## source

## Location/Qualifiers

1..868  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6059910"  
 /tissue\_type="large cell carcinoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 68"  
 /note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;  
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.8 kb. Library constructed by Life  
 Technologies."

## ORIGIN

Query Match 21.1%; Score 491; DB 5; Length 868;  
 Best Local Similarity 100.0%; Pred. No. 8.4e-252;  
 Matches 491; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 456 GAACAATTCAGTCCCTTCTCCATGAATGGTACACGAGAACATGTTGAGGAGAGAC 515  
 DB 11 GAACAATTCAGTCCCTTCTCCATGAATGGTACACGAGAACATGTTGAGGAGAGAC 70  
 QY 516 TTGCAACAAGAAGATGAGCCACAGAGAGAACACAAACAAGAGGATGATGTTCTTATG 575  
 DB 71 TTGCAACAAGAAGATGAGCCACAGAGAGAACACAAACAAGAGGATGATGTTCTTATG 130  
 QY 576 CGCACTGATGTAGATGATGATTTGAGACCTTGGAACTGAACTGATCTCATGAGAAC 635  
 DB 131 CGCACTGATGTAGATGATGATTTGAGACCTTGGAACTGAACTGATCTCATGAGAAC 190  
 QY 636 GACATGATTTACCACTGGAAGAGACAGTTTCAACAAGCTGTATCAGGATATGGAAGAG 695  
 DB 191 GAGCATGTTACCACTGGAAGAGACAGTTTCAACAAGCTGTATCAGGATATGGAAGAG 250  
 QY 696 ATGATGCTCAGCAGAGAAATCCAGATTCAGTGAACCACTAGTAGAAGATGAAAGATTG 755  
 DB 251 ATGATGCTCAGCAGAGAAATCCAGATTCAGTGAACCACTAGTAGAAGATGAAAGATTG 310  
 QY 756 CACCATGATACAGATGATTAACATACCAAGTCTATCAGGAAACAAGCTATATGAACCT 815  
 DB 311 CACCATGATACAGATGATTAACATACCAAGTCTATCAGGAAACAAGCTATATGAACCT 370  
 QY 816 CTAGAAATGAAGGGATGAAATCAGAGTACTGCTCCCTCAGGATATCTCTGA 875  
 DB 371 CTAGAAATGAAGGGATGAAATCAGAGTACTGCTCCCTCAGGATATCTCTGA 430  
 QY 876 GAGATTCACAGTAAATGTTAGAGAGAGTAAAGCAATTTTCTCTGGAAGAACAGCAGGAA 935  
 DB 431 GAGATTCACAGTAAATGTTAGAGAGAGTAAAGCAATTTTCTCTGGAAGAACAGCAGGAA 490  
 QY 936 GTACCAACACAG 946  
 DB 491 GTACCAACACAG 501

## RESULT 22

CN368829

## LOCUS

CN368829 1700600076582 GRN\_PFEHEP Homo sapiens cDNA 5', mRNA sequence.

ACCESSION CN368829

VERSION CN368829.1 GI:47368763

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 520)

Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,

Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,

Lebkowski, J. and Stanton, L.W.

Transcriptome characterization elucidates signaling networks that

control human ES cell growth and differentiation

## TITLE

JOURNAL  
COMMENT  
Nat. Biotechnol. 22 (6), 707-716 (2004)  
Contact: Brandenberger R  
Regenerative Medicine  
Geront Corporation  
230 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8658  
Fax: 650 473 7760  
Email: rbrandenberger@geron.com  
Insert Length: 520 Std Error: 0.00.

FEATURES  
source  
1. 520  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cells, DMSO-treated H9 cell line"  
/clone\_lib="GRN PREHEP"  
/note="oligo dt primed, full-length enriched cDNA library from DMSO-treated h9 cell line H9 (p22) maintained in feeder-free conditions"

ORIGIN  
Query Match 20.5%; Score 476; DB 7; Length 520;  
Best Local Similarity 100.0%; Pred. No. 9.1e-244;  
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 CCAGAAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTCTCCATGAAATGGTACA 490  
Db 7 CCAGAAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTCTCCATGAAATGGTACA 66  
QY 491 CGCAACACATGTTGAGGAGAGAGACTGCAACAAAGAGATGGACCCAGAGAGAACACCA 550  
Db 67 CGCAACACATGTTGAGGAGAGAGACTGCAACAAAGAGATGGACCCAGAGAGAACACCA 126  
QY 551 ACAAGAGATGATGATGTTCTTATGCGACTGATGATGATGATGATGATGATGATGATG 610  
Db 127 ACAAGAGATGATGATGTTCTTATGCGACTGATGATGATGATGATGATGATGATGATG 186  
QY 611 ACCTGAAGTATCTCATGAGAAACCGAGCATAGTTTACCAGTGGAGAGACAGATTTTACA 670  
Db 187 ACCTGAAGTATCTCATGAGAAACCGAGCATAGTTTACCAGTGGAGAGACAGATTTTACA 246  
QY 671 AGACTGTAATCAGGATATGGAAGAGATGATGTTCTGAGCAGGAGAAATCCAGATTCAGTGA 730  
Db 247 AGACTGTAATCAGGATATGGAAGAGATGATGTTCTGAGCAGGAGAAATCCAGATTCAGTGA 306  
QY 731 ACCAGTAGTAGAAGATGAAGATTTGACCATGATACAGATGATGATGATGATGATGATGATG 790  
Db 307 ACCAGTAGTAGAAGATGAAGATTTGACCATGATACAGATGATGATGATGATGATGATGATG 366  
QY 791 TGAGGAACAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCAGAGAGTAAC 850  
Db 367 TGAGGAACAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCAGAGAGTAAC 426  
QY 851 TGCTCCCTCCCTGAGGATTAATTCCTGTAGAGATTCACAGGTAATGTTAGAGAGTAA 906  
Db 427 TGCTCCCTCCCTGAGGATTAATTCCTGTAGAGATTCACAGGTAATGTTAGAGAGTAA 482

RESULT 23  
CN368823 770 bp mRNA linear EST 16-MAY-2004  
LOCUS 328775548 GRN\_ES Homo sapiens cDNA 5', mRNA sequence.  
DEFINITION CN368823  
ACCESSION CN368823.1 GI:47368757  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 770)  
REFERENCE  
AUTHORS  
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,  
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,

TITLE  
JOURNAL  
COMMENT  
Lebkowski, J and Stanton, L.W.  
Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation  
Nat. Biotechnol. 22 (6), 707-716 (2004)  
Contact: Brandenberger R  
Regenerative Medicine  
Geront Corporation  
230 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8658  
Fax: 650 473 7760  
Email: rbrandenberger@geron.com  
Insert Length: 770 Std Error: 0.00.

FEATURES  
source  
1. 770  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cells, cell lines H1, H7, and H9"  
/clone\_lib="GRN ES"  
/note="oligo dt primed, full-length enriched cDNA library from undifferentiated h9 cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"

ORIGIN  
Query Match 19.8%; Score 461; DB 7; Length 770;  
Best Local Similarity 99.8%; Pred. No. 1.2e-235;  
Matches 511; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGCACCGTCCATGCGCCAGCGTAAAGATGCCAAGAGCAGCGCGCAACAGCAGCAGCAGCG 60  
Db 186 CGCACCGTCCATGCGCCAGCGTAAAGATGCCAAGAGCAGCGCGCAACAGCAGCAGCAGCG 245  
QY 61 GCTCCGCGCAGCGTAGCAGAGTGC CGGCGAGCAGCGCGCGCGCGCGCGCGCGAGACAA 120  
Db 246 GCTCCGCGCAGCGTAGCAGAGTGC CGGCGAGCAGCGCGCGCGCGCGCGCGAGACAA 305  
QY 121 AGCATGAGAGACACAAGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
Db 306 AGCATGAGAGACACAAGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 365  
QY 181 GGTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240  
Db 366 GGTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 425  
QY 241 TTGTTGATGATGAGAGAGTCTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
Db 426 TTGTTGATGATGAGAGAGTCTAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 485  
QY 301 ATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
Db 486 ATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 545  
QY 361 CAGCAGTCCCGCCAG 420  
Db 546 CAGCAGTCCCGCCAG 605  
QY 421 AGGCAG 480  
Db 606 AGGCAG 665  
QY 481 AAATGGTACACGACAG 512  
Db 666 AAATGGTACACGACAG 697

RESULT 24  
BG720220 689 bp mRNA linear EST 08-MAY-2001  
LOCUS 602692270F1 NIH\_MGC\_97 Homo sapiens cDNA clone IMAGE:4824712 5',  
DEFINITION mRNA sequence.  
ACCESSION BG720220  
VERSION BG720220.1 GI:13999407



Db 301 TTTCCCAATGATATCTTCCTTAAATAATGACCTTGGGTGGGATACCTCTTGATAGAGAT 360

Qy 1419 AATGACAATCAAGAAGTTTATGAGAGGTGCTGAGTGTGACACCTAATGATGCTTT 1478

Db 361 AATGACAATCAAGAAGTTTATGAGAGGTGCTGAGTGTGACACCTAATGATGCTTT 420

Qy 1479 GCTAAATGCTAATGATGCTTCACTCCTGAAGGCACA 1513

Db 421 GCTAAATGCTAATGATGCTTCACTCCTGAAGGCACA 455

RESULT 26

BP328261 583 bp mRNA linear EST 17-SEP-2004

LOCUS BP328261 Sugano cDNA library, rectum Homo sapiens cDNA clone

DEFINITION RCT00312, mRNA sequence.

ACCESSION BP328261.1 GI:52257341

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 583)

AUTHORS Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J., Mizushima-Sugano, J., Nakai, K. and Sugano, S.

TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions

JOURNAL Genome Res. 14 (9), 1711-1718 (2004)

COMMENT Contact: Yutaka Suzuki Department of Virology Institute of Medical Science, University of Tokyo 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan Email: yusuzuki@ms.u-tokyo.ac.jp

FEATURES

source 1..583

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="RCT00312"

/tissue\_type="rectum"

/clone\_lib="Sugano cDNA library, rectum"

ORIGIN

Query Match 19.4%; Score 452; DB 5; Length 583;

Best Local Similarity 99.6%; Pred. No. 7.6e-231; Indels 0; Gaps 0;

Matches 552; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1366 ATGATACTTCCCTTAAATAATGACCTTGGCTGGGATACCTCTTGATAGAGATATGACA 1425

Db 15 ATGATACTTCCCTTAAATAATGACCTTGGCTGGGATACCTCTTGATAGAGATATGACA 74

Qy 1426 ATGCAAGAAGTTTATGAAGGTGCTGAGTGTGACACCTAATGATGCTTGTGTAAG 1485

Db 75 ATGCAAGAAGTTTATGAAGGTGCTGAGTGTGACACCTAATGATGCTTGTGTAAG 134

Qy 1486 TCCATTATGCTTCATCTCTGAGGACACAAATTCCTGAGAGCATCCCATATTAA 1545

Db 135 TCCATTATGCTTCATCTCTGAGGACACAAATTCCTGAGAGCATCCCATATTAA 194

Qy 1546 AGGAAGGAATAGAAATCCGGAGATCCTTGGCACTGATGATGGGAGATTTTATTTCCACTGG 1605

Db 195 AGGAAGGAATAGAAATCCGGAGATCCTTGGCACTGATGATGGGAGATTTTATTTCCACTGG 254

Qy 1606 GGGATGCCATGACAGAGGTTGGCAACAAAGAGGCATATAAGTGGTATGAGCTTTGGGCACA 1665

Db 255 GGGATGCCATGACAGAGGTTGGCAACAAAGAGGCATATAAGTGGTATGAGCTTTGGGCACA 314

Qy 1666 AGAGAGGACACTTTGATCTGTCTGGCAACCGCTCACTCTACATGTGAATGGACTGAAG 1725

Db 315 AGAGAGGACACTTTGATCTGTCTGGCAACCGCTCACTCTACATGTGAATGGACTGAAG 374

Qy 1726 CACAGCCTTGGTGGACCCCAAAAGAAACGGGCTACAGAGTTAGTAAGTCTTTAGAAA 1785

Db 375 CACAGCCTTGGTGGACCCCAAAAGAAACGGGCTACAGAGTTAGTAAGTCTTTAGAAA 434

Qy 1786 GAACTGGAAGTTAATCCGAGATGAAGCCTTGGAGTATGATGATGAACCAAGGTCTCT 1845

Db 435 GAACTGGAAGTTAATCCGAGATGAAGCCTTGGAGTATGATGATGAACCAAGGTCTCT 494

Qy 1846 TCCTGCTGAGGATGAAAACCTTGAAGGAAAAGGGGACTTGGAGCCAGTTCAAGTGTGGC 1905

Db 495 TCCTGCTGAGGATGAAAACCTTGAAGGAAAAGGGGACTTGGAGCCAGTTCAAGTGTGGC 554

Qy 1906 AGCAAGGAAGAAGA 1919

Db 555 AGCAAGGAAGAAGA 568

RESULT 27

BQ889709 857 bp mRNA linear EST 16-AUG-2002

LOCUS BQ889709

DEFINITION AGENCOURT 8118183 Lupski dorsal root ganglion Homo sapiens cDNA clone IMAGE:6181443 5', mRNA sequence.

ACCESSION BQ889709.1 GI:22281723

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 857)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D. Email: cgaops-x@mail.nih.gov Tissue Procurement: Dr. James R. Lupski CDNA Library Preparation: Life Technologies, Inc. DNA sequencing by: The I.M.A.G.E. Consortium (LNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov Plate: LLM13565 Row: n Column: 04 High quality sequence stop: 590.

FEATURES

Location/Qualifiers

1..857

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:6181443"

/sex="male"

/tissue\_type="dorsal root ganglia"

/dev\_stage="adult, 36 yr"

/lab\_host="DH10B"

/clone\_lib="Lupski dorsal root ganglion"

/note=vector: pCMV-SPORT6 (Life Technologies); Site\_1: NotI; Site\_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCACCGCTCCG-3' and 5'-GACTAGTCTTAGTCGAGCGGCGGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

ORIGIN

Query Match 19.4%; Score 450; DB 5; Length 857;

Best Local Similarity 99.8%; Pred. No. 9.8e-230; Indels 0; Gaps 0;

Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGACCTGCAATGGCCAGCGTGAAGATGCCAAGAGCGGCAACAGCAGCAGCG 60



CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLCM2139 row: i column: 04  
High quality sequence stop: 595.  
Location/Qualifiers  
1. .1014  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5952075"  
/tissue\_type="ductal carcinoma, cell line"  
/lab\_host="NIH\_MGC\_110"  
/clone\_lib="NIH\_MGC\_110"  
/note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

ORIGIN  
Query Match 19.4%; Score 450; DB 5; Length 1014;  
Best Local Similarity 99.8%; Pred. No. 1e-229;  
Matches 500; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGGACCGTGCATGGCCAGCGTGAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
DB 103 CGGACCGTGCATGGCCAGCGTGAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 162  
QY 61 GCTCCGCGCAGCGGTAGCAGCGTGGCGCAGCAGCGCCCGGGCCCGGAGAGAGACAA 120  
DB 163 GCTCCGCGCAGCGGTAGCAGCGTGGCGCAGCAGCGCCCGGGCCCGGAGAGAGACAA 222  
QY 121 AGCATGGAGGACACAGAAATGGGAGAAAGCGGACTCTCGGGAACCTTCACTTTCACGT 180  
DB 223 AGCATGGAGGACACAGAAATGGGAGAAAGCGGACTCTCGGGAACCTTCACTTTCACGT 282  
QY 181 GGTATTATGTGATTCGATTGCTGGCGCTGTGACATCTGTAGCTGTCTGTTGGTTGATC 240  
DB 283 GGTATTATGTGATTCGATTGCTGGCGCTGTGACATCTGTAGCTGTCTGTTGGTTGATC 342  
QY 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGTTGATGAG 300  
DB 343 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGTTGATGAG 402  
QY 301 ATTTTGATGCTGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 360  
DB 403 ATTTTGATGCTGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 462  
QY 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGAGAGAGGTTCTCTGAG 420  
DB 463 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGAGAGAGGTTCTCTGAG 522  
QY 421 AGGCAGAACCCCGAGATATCGAGATGAGCAAGCAAAATTCAGTCCCTTCTCCATG 480  
DB 523 AGGCAGAACCCCGAGATATCGAGATGAGCAAGCAAAATTCAGTCCCTTCTCCATG 582  
QY 481 AAATGGTACACGAGAACATG 501  
DB 583 AAATGGTACACGAGAACATG 603

RESULT 30  
AUI32666  
LOCUS  
DEFINITION  
AUI32666 NT2RP4 Homo sapiens cDNA clone NT2RP4000310 5', mRNA  
sequence.  
AUI32666  
ACCESSION

AUI32666.1 GI:10993205  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 767)  
Ota.T., Sugiyama.T., Ishii.S., Suzuki.Y., Saito.K., Yamamoto.J., Nishikawa.T., Nakamura.Y., Nagai.T., Sugano.S., Masuho.Y. and Isogai.T.  
HRI human cDNA project (Ota.T., Sugiyama.T., Ishii.S., Suzuki.Y., Saito.K., Yamamoto.J., Nishikawa.T., Nakamura.Y., Nagai.T., Sugano.S., Masuho.Y., Isogai.T.)  
Unpublished (2000)  
Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@hri.co.jp  
HRI human cDNA project; 5' & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.  
Location/Qualifiers  
1. .767  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="NT2RP4000310"  
/cell\_type="teratocarcinoma"  
/clone\_lib="NT2"  
/notes="Vector: pME18SFL3; mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"  
Query Match 19.3%; Score 448; DB 1; Length 767;  
Best Local Similarity 99.6%; Pred. No. 1.e-228;  
Matches 548; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CGGACCGTGCATGGCCAGCGTGAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
DB 181 CGGACCGTGCATGGCCAGCGTGAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 240  
QY 61 GCTCCGCGCAGCGGTAGCAGCGTGGCGCAGCAGCGCCCGGGCCCGGAGAGAGACAA 120  
DB 241 GCTCCGCGCAGCGGTAGCAGCGTGGCGCAGCAGCGCCCGGGCCCGGAGAGAGACAA 300  
QY 121 AGCATGGAGGACACAGAAATGGGAGAAAGCGGACTCTCGGGAACCTTCACTTTCACGT 180  
DB 301 AGCATGGAGGACACAGAAATGGGAGAAAGCGGACTCTCGGGAACCTTCACTTTCACGT 360  
QY 181 GGTATTATGTGATTCGATTGCTGGCGCTGTGACATCTGTAGCTGTCTGTTGGTTGATC 240  
DB 361 GGTATTATGTGATTCGATTGCTGGCGCTGTGACATCTGTAGCTGTCTGTTGGTTGATC 420  
QY 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGTTGATGAG 300  
DB 421 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGTTGATGAG 480  
QY 301 ATTTTGATGCTGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 360  
DB 481 ATTTTGATGCTGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 540  
QY 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGAGAGAGGTTCTCTGAG 420  
DB 541 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGAGAGAGGTTCTCTGAG 600  
QY 421 AGGCAGAACCCCGAGATATCGAGATGAGCAAGCAAAATTCAGTCCCTTCTCCATG 480  
DB 601 AGGCAGAACCCCGAGATATCGAGATGAGCAAGCAAAATTCAGTCCCTTCTCCATG 660

VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT







**TITLE** Transcriptional characterization elucidates signaling networks that control human ES cell growth and differentiation

**JOURNAL** Nat. Biotechnol. 22 (6), 707-716 (2004)

**COMMENT** Contact: Brandenberger R  
Regenerative Medicine  
Geron Corporation  
230 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8658  
Fax: 650 473 7760  
Email: rbrandenberger@geron.com  
Insert Length: 695 Std Error: 0.00.  
Location/Qualifiers  
1..695  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cells, embryoid bodies  
derived from H1, H7 and H9 cells"  
/clone\_lib="GRN\_EB"  
/note="oligo dT primed, full-length enriched cDNA library  
from embryoid body outgrowths derived from hES cell lines  
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free  
conditions."

**FEATURES**

**source**  
1..695  
Query Match 18.5%; Score 430; DB 7; Length 695;  
Best Local Similarity 99.8%; Pred. No. 5.4e-219;  
Matches 480; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

**ORIGIN**

1 CGGACCGTGCATCGCCAGCGTAAGAATGCCAAGCAGCGGCAACAGCAGCAGCGG 60  
215 CGGACCGTGCATCGCCAGCGTAAGAATGCCAAGCAGCGGCAACAGCAGCAGCGG 274  
61 GCTCCGGCAGCGGTAGCAGCGTCCGCGCAGCAGCGCCCGCGGCCCGGAGAGACAA 120  
275 GCTCCGGCAGCGGTAGCAGCGTCCGCGCAGCAGCGCCCGCGGCCCGGAGAGACAA 334  
121 AGCATGGAGGACACAGAAATGGAGAAAGCGGCACTCTCGGAACTTCATCTTCACGT 180  
335 AGCATGGAGGACACAGAAATGGAGAAAGCGGCACTCTCGGAACTTCATCTTCACGT 394  
181 GGTTCATGGTATTCATTCGCTGGCGCTCTGGACATCTGTAGCTGTCTGTTGTTGATC 240  
395 GGTTCATGGTATTCATTCGCTGGCGCTCTGGACATCTGTAGCTGTCTGTTGTTGATC 454  
241 TTGTTGACTATGAGGAATTCATAGGAATCTAGGAATCTATGATGCTGTGATGGAG 300  
455 TTGTTGACTATGAGGAATTCATAGGAATCTAGGAATCTATGATGCTGTGATGGAG 514  
301 ATTTTGATGGATGATGCCAAGTTTATTTAGGACTTAAGAGAGATCTACTTCAGAGC 360  
515 ATTTTGATGGATGATGCCAAGTTTATTTAGGACTTAAGAGAGATCTACTTCAGAGC 574  
361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACTGAGCCCGCAGGAGCGTTCTCTGTGG 420  
575 CAGCAGTCCCGCAGAGAGGCTGAGCCACACTGAGCCCGCAGGAGCGTTCTCTGTGG 634  
421 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAGAAACAATTCAGTCCCTTCTCCATG 480  
635 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAGAAACAATTCAGTCCCTTCTCCATG 694  
481 A 481  
695 A 695

**RESULT 35**

AL700393  
LOCUS DKEZp686K01118 r1 686 554 bp mRNA linear EST 04-SEP-2003  
DEFINITION DKEZp686K01118 5', mRNA sequence.  
ACCESSION AL700393  
VERSION AL700393.1 GI:19620926

**KEYWORDS**

Homo sapiens (human)

**SOURCE****ORGANISM****REFERENCE****AUTHORS****TITLE****JOURNAL****COMMENT****EST.**

Homo sapiens

Homo sapiens

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

Pousetka, A., Wellenreuther, R., Mewes, H.W., Weil, B. and Wiemann, S.)

EST (Pousetka, A., Wellenreuther, R., Mewes, H.W., Weil, B. and Wiemann, S.)

Unpublished (1999)

Contact: MIPS

MIPS

Ingolstaedter Landstr. 1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;

Sequenced by DKFZ (German Cancer Research Center,

Heidelberg/Germany) within the cDNA sequencing consortium of the

German Genome Project.

No 5' sequence available.

This clone (DKFZp686K01118) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

Location/Qualifiers

1..554

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="DKFZp686K01118"

/dev\_stage="adult"

/lab\_host="DH10B"

/clone\_lib="686 (synonym: hlcc3)"

/notes="Vector: pTriplex2; Site 1: SfiIA; Site 2: SfiIB;  
cDNA-collection"**ORIGIN**

Query Match 18.3%; Score 425; DB 1; Length 554;

Best Local Similarity 99.8%; Pred. No. 2.6e-216;

Matches 475; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1439 TTATGAAGAGGTCTCAGTGTGACACCTTAATGATGGCTTTGCTAAAGTCCATTATGGCTT 1498

Db 1 TTATGAAGAGGTCTCAGTGTGACACCTTAATGATGGCTTTGCTAAAGTCCATTATGGCTT 60

QY 1499 CATCTCAAGCAGACAGAAATTCCTGAGAGCATCCCATATTTAAAGAGGAATAGA 1558

Db 61 CATCTCAAGCAGACAGAAATTCCTGAGAGCATCCCATATTTAAAGAGGAATAGA 120

QY 1559 ATCCGAGATCTTGGCACTGATGGAGATTTTATTTCCACCTGGGGGATGCCATGCA 1618

Db 121 ATCCGAGATCTTGGCACTGATGGAGATTTTATTTCCACCTGGGGGATGCCATGCA 180

QY 1619 GAGGTTTGGGAAACAAAGAGGCATATAAGTGTATGAGCTTTGGGCAACAGAGGACACTT 1678

Db 181 GAGGTTTGGGAAACAAAGAGGCATATAAGTGTATGAGCTTTGGGCAACAGAGGACACTT 240

QY 1679 TGCACTCTGTGCGCAACGCTCACTCAATGATGAGTGAATGAGTGAATGAGTGAATGAGTGA 1738

Db 241 TGCACTCTGTGCGCAACGCTCACTCAATGATGAGTGAATGAGTGAATGAGTGAATGAGTGA 300

QY 1739 GACCCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTTAAAGAAAGAACTGAAGTT 1798

Db 301 GACCCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTTAAAGAAAGAACTGAAGTT 360

QY 1799 AATCCGAGATGAAGGCTTTGCACTGATGATGAATAAGCCAAAGGCTCTCTCTCCCTGAGGA 1858

Db 361 AATCCGAGATGAAGGCTTTGCACTGATGATGAATAAGCCAAAGGCTCTCTCTCCCTGAGGA 420

QY 1859 TGAACACCTGAGGAAAAAGGGGACTGAGCCAGTTTACGCTGTGTCAGCAGGAA 1914

Db 421 TGAACACCTGAGGAAAAAGGGGACTGAGCCAGTTTACGCTGTGTCAGCAGGAA 476



```

Db      61  GTTTTATTAGGACTTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCGCCAGAGAGGCT 120
Qy      384  GAGCCACACACTGAGCCGAGGAGCAGGTTCTGTGGAGGAGAGACCCAGCAATATCGAA 443
Db      121  GAGCCACACACTGAGCCGAGGAGCAGGTTCTGTGGAGGAGAGACCCAGCAATATCGAA 180
Qy      444  GATGAGCAAAAGAACAAATTCAGTCCCTTCTCCATGAATGTTACAGCAGACATGTT 503
Db      181  GATGAGCAAAAGAACAAATTCAGTCCCTTCTCCATGAATGTTACAGCAGACATGTT 240
Qy      504  GAGGAGAGAGACTTGCACAGAGATGAGCCACAGAGAGAACCAACAGAGAGATGAT 563
Db      241  GAGGAGAGAGACTTGCACAGAGATGAGCCACAGAGAGAACCAACAGAGAGATGAT 300
Qy      564  GAGTTTCTTATGGCGACTGATGTAGATGATAGATTGAGACCCCTGGAACTGAGTATCT 623
Db      301  GAGTTTCTTATGGCGACTGATGTAGATGATAGATTGAGACCCCTGGAACTGAGTATCT 360
Qy      624  CATGAGAAACCGAGCATAGTTACCGTGGAGAGACAGAGTTTACAGAGACTGTAATCAG 683
Db      361  CATGAGAAACCGAGCATAGTTACCGTGGAGAGACAGAGTTTACAGAGACTGTAATCAG 420
Qy      684  GATA 687
Db      421  GATA 424

```

## RESULT 38

```

CN481470
LOCUS      606 bp mRNA linear EST 26-APR-2004
DEFINITION hw08b05.y1 Human primary human ocular pericytes. Unamplified (hw)
ACCESSION  CN481470
VERSION     CN481470.1 GI:46562974
SOURCE      Homo sapiens (human)

```

## ORGANISM

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Tsai, J.Y. and Wistow, G.

```

Expressed sequence tag analysis of cultured primary human ocular pericytes

## JOURNAL

```

Unpublished (2004)
Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov

```

Plate: 08 row: b column: 05

Seq primer: M13RP1 reverse primer (ABI).

Location/Qualifiers

## FEATURES

## source

```

1..606
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="hw08b05"
/cell_type="pericytes"
/dev_stage="Adult"
/lab_host="EMDH108"
/clone_lib="Human primary human ocular pericytes.
Unamplified (hw)"
/note="Organ: Eye; Vector: pSport1; RNA was extracted from
primary human pericytes in culture. A directionally cloned
cDNA library in the pSPOR1 vector (Invitrogen) was
constructed at Bioserve Biotechnology (Laurel MD)
essentially following the protocols of the SuperScript
Plasmid System full details of which are contained in the
manufacturer's instruction manual
(http://www.lifetech.com/). First strand synthesis was

```

carried out using a Not I primer-adaptor  
[5'-pGATAGTCTAGATCGAGCGCCGCCC(T)15-3']. cDNA was  
cloned in Not I/Sal I sites. EST analysis was performed at  
the NIH Intramural Sequencing Center (NISC)."

## ORIGIN

```

Query Match      18.2%; Score 424; DB 7; Length 606;
Best Local Similarity 99.8%; Pred. No. 9e-216;
Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  CGGACCGGTGCAATGGCCCGAGCGTAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCG 60
Db      132  CGGACCGGTGCAATGGCCCGAGCGTAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCG 191
Qy      61  GCTCCGGCAGCGGTAGCAGCAGTGGCGGCGAGCGCGCGCGCGCGCGCGCGCGAGACAA 120
Db      192  GCTCCGGCAGCGGTAGCAGCAGTGGCGGCGAGCGCGCGCGCGCGCGCGCGAGACAA 251
Qy      121  AGCATGGAGGACACAGAGATGGAGAAAGCGGACTCTCGGGAATTCATTCTTCACGT 180
Db      252  AGCATGGAGGACACAGAGATGGAGAAAGCGGACTCTCGGGAATTCATTCTTCACGT 311
Qy      181  GGTATTATGGTATGCAATTGCTGGCGCTCTGACATCTGTAGCTGTGTTGTTGATC 240
Db      312  GGTATTATGGTATGCAATTGCTGGCGCTCTGACATCTGTAGCTGTGTTGTTGATC 371
Qy      241  TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATCTGATGTTGATGAG 300
Db      372  TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATCTGATGTTGATGAG 431
Qy      301  ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db      432  ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 491
Qy      361  CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGAGGTTCTGTGG 420
Db      492  CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGAGGTTCTGTGG 551
Qy      421  AGGCAGAACCCAGATATCGAAGATGAAGCAAAAGCAAAATTCAGTCCCTTCT 475
Db      552  AGGCAGAACCCAGATATCGAAGATGAAGCAAAAGCAAAATTCAGTCCCTTCT 606

```

## RESULT 39

```

BF943222
LOCUS      504 bp mRNA linear EST 22-JAN-2001
DEFINITION QV2-NN0045-041000-403-b08 NN0045 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BF943222
VERSION     BF943222.1 GI:12360497
KEYWORDS    EST.
SOURCE      Homo sapiens (human)

```

## ORGANISM

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 504)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

```

## REFERENCE

## AUTHORS

Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

## JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

## MEDLINE

20202663

## PUBMED

10737800

## COMMENT

Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV2&t2=QV2-NN0045-041000-403-b08&t3=2000-10-04&t4=1)  
041000-403-b08&t3=2000-10-04&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 504.

FEATURES  
Location/Qualifiers  
1..504  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="NN0045"  
/note="Organ: nervous normal; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196, 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN  
Query Match 18.2%; Score 423; DB 4; Length 504;  
Best Local Similarity 99.8%; Pred. No. 3e-215;  
Matches 473; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 332 AGGACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACA 391  
Db 31 AGGACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACA 90  
QY 392 CACTGAGCCGAGGAGCAGTCTCTGTGGAGCGAGACCCAGAAATATCGAAGATGAAGC 451  
Db 91 CACTGAGCCGAGGAGCAGTCTCTGTGGAGCGAGACCCAGAAATATCGAAGATGAAGC 150  
QY 452 AAAGCAACAATTCAGTCCCTTCTCCATCAATGTGTACACGACAGAAATGTGTAGGAGGA 511  
Db 151 AAAGCAACAATTCAGTCCCTTCTCCATCAATGTGTACACGACAGAAATGTGTAGGAGGA 210  
QY 512 AGACTTGCNACAAAGAGATGGACCCAGGAGAACCAACAGAGGATGATGATTTCT 571  
Db 211 AGACTTGCNACAAAGAGATGGACCCAGGAGAACCAACAGAGGATGATGATTTCT 270  
QY 572 TATGGCGACTGATGTAGATGATAGATTTCAGACCCCTGGAACTGAACTATCTCATGAAGA 631  
Db 271 TATGGCGACTGATGTAGATGATAGATTTCAGACCCCTGGAACTGAACTATCTCATGAAGA 330  
QY 632 AACCCGACATAGTTACCCAGTGGAGAGACAGTTTCAAGACTGTATCAGGATATGGA 691  
Db 331 AACCCGACATAGTTACCCAGTGGAGAGACAGTTTCAAGACTGTATCAGGATATGGA 390  
QY 692 AGAGATGATGTCTGAGCAGGAAATCCAGATTCCAGTGAAACAGTAGTAGAAGATGAAG 751  
Db 391 AGAGATGATGTCTGAGCAGGAAATCCAGATTCCAGTGAAACAGTAGTAGAAGATGAAG 450  
QY 752 ATTGCACCATGATACAGATGATGTATACATACCAAGTCTATGAGGAACAGCAGT 805  
Db 451 ATTGCACCATGATACAGATGATGTATACATACCAAGTCTATGAGGAACAGCAGT 504

RESULT 40  
CN481475  
LOCUS  
DEFINITION  
hw08b10.y1 Human primary human ocular pericytes. Unamplified (hw)  
Homo sapiens cDNA clone hw08b10 5', mRNA sequence.  
CN481475  
ACCESSION  
VERSION  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 610)

AUTHORS  
TITLE  
JOURNAL  
COMMENT

Tsai, J.Y. and Wistow, G.  
Expressed sequence tag analysis of cultured primary human ocular pericytes  
Unpublished (2004)  
Contact: Wistow G  
Section on Molecular Structure and Function  
National Eye Institute  
6/331, NIH, Bethesda, MD 20892-2740, USA  
Tel: 301 402 3452  
Fax: 301 496 0078  
Email: graeme@helix.nih.gov  
Plate: 08 row: b column: 10  
Seq primer: M13RP1 reverse primer (ABI).  
Location/Qualifiers  
1..610  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="hw08b10"  
/cell\_type="pericytes"  
/dev\_stage="Adult"  
/lab\_host="EMDH10B"  
/clone\_lib="Human primary human ocular pericytes."  
Unamplified (hw)  
/note="Organ: Eye; Vector: pSport1; RNA was extracted from primary human pericytes in culture. A directionally cloned cDNA library in the pSPORT1 vector (Invitrogen) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGACTAGTCTAGATCGGCGCCGCTT)15-3']. cDNA was cloned in Not I/Sal I sites. EST analysis was performed at the NIH Intramural Sequencing Center (NISC)."

ORIGIN

Query Match 18.2%; Score 423; DB 7; Length 610;  
Best Local Similarity 99.8%; Pred. No. 3.1e-215;  
Matches 473; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGGACCTGCAATGGCCGCGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
Db 137 CGGACCTGCAATGGCCGCGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 196  
QY 61 GCTCCGCGCAGCGGTAGCACGAGTCCGGGCGAGCAGCGCCCGGGCGCCGAGAGAGACAA 120  
Db 197 GCTCCGCGCAGCGGTAGCACGAGTCCGGGCGAGCAGCGCCCGGGCGCCGAGAGAGACAA 256  
QY 121 AGCATGAGGACACAGAAATGGGAGGAAAGCGGACTCTCGGAACTTCACTTCTCAGCT 180  
Db 257 AGCATGAGGACACAGAAATGGGAGGAAAGCGGACTCTCAGGAACTTCACTTCTCAGCT 316  
QY 181 GGTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240  
Db 317 GGTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 376  
QY 241 TTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300  
Db 377 TTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 436  
QY 301 ATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
Db 437 ATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 496  
QY 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGGATTCCTGTGG 420  
Db 497 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGGATTCCTGTGG 556  
QY 421 AGGAGAGACCCCGAGAGATTCGAGATGAGCAAGAGACAAATTCAGTCCCTTC 474  
Db 557 AGGAGAGACCCCGAGAGATTCGAGATGAGCAAGAGACAAATTCAGTCCCTTC 610

```
RESULT 41
AUI38660
LOCUS
DEFINITION AUI38660 PLACE1 Homo sapiens cDNA clone PLACE1009036 5', mRNA
ACCESSION AUI38660
VERSION AUI38660
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 790)
Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Ishigai,T.
HRI human cDNA project
Unpublished (2000)
CONTACT: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
Location/Qualifiers
source
1..790
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1009036"
/tissue_type="placenta"
/clone_lib="PLACE1"
/note="Vector: pME18SFL3"
ORIGIN
Query Match 18.1%; Score 420; DB 1; Length 790;
Best Local Similarity 100.0%; Pred. NO. 1.3e-213;
Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 749 AAGATTGCACCATGATACAGATGATGTAACATACCAAGTCTATGAGGAAACAAGCAGTATA 808
DB 1 AAGATTGCACCATGATACAGATGATGTAACATACCAAGTCTATGAGGAAACAAGCAGTATA 60
QY 809 TGAACCTCTAGAAAATGAAGGGATAGAAATACAGAGTAACTGCTCCCTGAGGATAA 868
DB 61 TGAACCTCTAGAAAATGAAGGGATAGAAATACAGAGTAACTGCTCCCTGAGGATAA 120
QY 869 TCCTGTAGAGATTCACAGTAAATGTTAGAGAGTAAAGCATTTTCTGTGGAGAGACA 928
DB 121 TCCTGTAGAGATTCACAGTAAATGTTAGAGAGTAAAGCATTTTCTGTGGAGAGACA 180
QY 929 GCAGAGATTCACAGAGAAACAAATAGAAAAACAGATGATCCAGAGCAAAAAAGT 988
DB 181 GCAGAGATTCACAGAGAAACAAATAGAAAAACAGATGATCCAGAGCAAAAAAGT 240
QY 989 TAAGAAAAAGAGCGCTAAACTTTTAAATTAATTTGATAAGCATTTAAAGCTGAACCTGA 1048
DB 241 TAAGAAAAAGAGCGCTAAACTTTTAAATTAATTTGATAAGCATTTAAAGCTGAACCTGA 300
QY 1049 TGCTGCAGAAAAACTCCGTAAAGGGGAAAAATTTGAGAGAGCGAGTGAATGCAATTTAAAGA 1108
DB 301 TGCTGCAGAAAAACTCCGTAAAGGGGAAAAATTTGAGAGAGCGAGTGAATGCAATTTAAAGA 360
QY 1109 ACTAGTAGCGAAATACCTCTCAGAGTCCACGAGCAAGATATGGGAAGCGCAGTGTGAGGA 1168
DB 361 ACTAGTAGCGAAATACCTCTCAGAGTCCACGAGCAAGATATGGGAAGCGCAGTGTGAGGA 420

RESULT 42
BQ318605
LOCUS
DEFINITION BQ318605 CT0549-210900-561-a04 CT0549 Homo sapiens cDNA, mRNA sequence.
ACCESSION BQ318605
VERSION BQ318605.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 433)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
PUBLISHED
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM0&t2=CM0-CT0549-
210900-561-a04&t3=2000-09-21&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 433.
Location/Qualifiers
source
1..433
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0549"
/note="Organ: colon; Vector: puc18; Site_1: Sma1; Site_2:
Sma1; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
ORIGIN
Query Match 18.0%; Score 418; DB 5; Length 433;
Best Local Similarity 100.0%; Pred. No. 1.4e-212;
Matches 418; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 479 TGAATGTTACACGACGACACATTTGAGGGAGAGACTTCACACAGAGATGGACCCAC 538
DB 16 TGAATGTTACACGACGACACATTTGAGGGAGAGACTTCACACAGAGATGGACCCAC 75
QY 539 AGAGAACCCACACACAGAGGATGATGATTTCTTATGGCGACTGATGATGATGATGAT 598
DB 76 AGAGAACCCACACACAGAGGATGATGATTTCTTATGGCGACTGATGATGATGATGAT 135
QY 599 TGAGACCCCTGGAACTCTGAAGTATCTCATGAAGAAACCGAGCATAGTACACGCTGGAGA 658
DB 136 TGAGACCCCTGGAACTCTGAAGTATCTCATGAAGAAACCGAGCATAGTACACGCTGGAGA 195
QY 659 GACAGTTTCCAAAGACTCTAATCAGGATATGGAAGAGATGATCTGAGCAGGAATCC 718
```



Db 196 GACAGTTTACAGACTGTGTAATCAGGATATGGAAGAGATGATCTCTGAGCAGCAAAATCC 255

QY 719 AGATTCCAGTGAACCACTAGTAGAAGATGAAGATGACCATGATACAGATGATGTAA 778

Db 256 AGATTCCAGTGAACCACTAGTAGAAGATGAAGATGACCATGATACAGATGATGTAA 315

QY 779 ATACCAAGTCTATGAGAAACAAGCAGTATATGACCTCTAGAAAATGAAGGGATAGAAAT 838

Db 316 ATACCAAGTCTATGAGAAACAAGCAGTATATGACCTCTAGAAAATGAAGGGATAGAAAT 375

QY 839 CACAGAAGTAACTGCTCCCTCCTGAGGATAATCTCTGTAAGAAGATTCACAGGTAATGTGA 896

Db 376 CACAGAAGTAACTGCTCCCTCCTGAGGATAATCTCTGTAAGAAGATTCACAGGTAATGTGA 433

RESULT 43

BQ278505 1100 bp mRNA linear EST 07-MAY-2002

LOCUS AGENCOURT\_6773215 NIH\_MGC\_128 Homo sapiens cDNA clone IMAGE:5811567

DEFINITION 5', mRNA sequence.

ACCESSION BQ278505

VERSION BQ278505.1 GI:20488713

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1100)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: NCI

cDNA Library Preparation: Michael Brownstein Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Cloning Strategy: Agencourt Bioscience Corporation

Cloning Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLCW2060 row: j column: 16

High quality sequence stop: 394.

Location/Qualifiers

1..1100

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5811567"

/tissue\_type="mixed (pool of 40 RNAs)"

/lab\_host="DH10B (T1-phage-resistant)"

/clone\_lib="NIH MGC 128"

/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc); Site 2: SfiI (ggccctctggcc); Double-stranded cDNA was prepared from a pool of 40 cell line polyA+ RNAs (bladder - 2%, blood - 33.4%, brain 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows:

5'-AGCAGTGTATACAGCAGATGGCATTCAGCGGG-3' and

5'-ATCTAGAGCCGAGCGCCACATG-dT(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the >2 kb size fraction (other fractions present in NIH MGC\_126 and NIH MGC\_127). Library created in the laboratory of T. Ueda, W.D., Ph.D. (NIMH, NIH). Note: this is a NIH\_MGC Library."

FEATURES

source

ORIGIN

Query Match 17.9%; Score 417; DB 5; Length 1100;

Best Local Similarity 100.0%; Pred. No. 5.9e-212;

Matches 417. Conservative 0; Mismatches 0; Indels 0; Gaps 0;

/clone\_lib="NN0045"  
/note="Organ: nervous normal; Vector: puc18; Site 1: Sma1;  
Site 2: Sma1; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

## ORIGIN

Query Match 17.9%; Score 415; DB 4; Length 444;  
Best Local Similarity 100.0%; Pred. No. 5.9e-211;  
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 332 AGACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACA 391  
DB 30 AGACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACA 89  
QY 392 CACTGAGCCCGAGGAGCAGGTTCTGTGGAGGAGAGCCAGATATCGAAGATGAAGC 451  
DB 90 CACTGAGCCCGAGGAGCAGGTTCTGTGGAGGAGAGCCAGATATCGAAGATGAAGC 149  
QY 452 AAAAGAGACAAATTCAGTCCCTCTCCATGAAATGGTACACGAGAGACATGTTGAGGAGA 511  
DB 150 AAAGAGACAAATTCAGTCCCTCTCCATGAAATGGTACACGAGAGACATGTTGAGGAGA 209  
QY 512 AGACTTGCACAGAGAGATGAGCCAGAGAGAACCAACAGAGAGGATGATGATTTCT 571  
DB 210 AGACTTGCACAGAGAGATGAGCCAGAGAGAACCAACAGAGAGGATGATGATTTCT 269  
QY 572 TATGGCGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 631  
DB 270 TATGGCGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 329  
QY 632 AACCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 691  
DB 330 AACCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 389  
QY 692 AGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 746  
DB 390 AGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 444

## RESULT 45

BQ639558  
LOCUS  
DEFINITION  
hel7b04.y1 Human Retina cDNA (Un-normalized, unamplified): hd/he  
Homo sapiens cDNA clone hel7b04 5', mRNA sequence.  
BQ639558  
VERSION  
BQ639558.1 GI:21764017  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1. (bases 1 to 570)  
Wistow, G., Bernstein, S.L., Wyatt, M.K., Ray, S., Behal, A.,  
Touchman, J.W., Bouffard, G., Smith, D. and Peterson, K.  
Expressed sequence tag analysis of human retina for the NEIBank  
Project: Retbindin, an abundant, novel retinal cDNA and alternative  
splicing of other retina-preferred gene transcripts  
Mol. Vis. 8 (4), 196-204 (2002)  
JOURNAL  
MEDLINE  
PUBMED  
22103461  
COMMENT  
Contact: Wistow G  
Section on Molecular Structure and Function  
National Eye Institute  
6/331, NIH, Bethesda, MD 20892-2740, USA  
Tel: 301 402 3452  
Fax: 301 496 0078  
Email: graeme@helix.nih.gov  
Plate: 17 row: b column: 04  
Seq primer: M13RP1 reverse primer (ABI).

FEATURES  
source

Location/Qualifiers  
1..570  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="hel7b04"  
/tissue\_type="Retina"  
/dev\_stage="Adult"  
/lab\_host="EMDH10B"  
/clone\_lib="Human Retina cDNA (Un-normalized,  
unamplified): hd/he"  
/note="Organ: Eye; Vector: pSPORT1; Neural retina tissue  
was dissected from two 80 year old donors with no observed  
eye disease. 100ug of total RNA was used for library  
construction. A directionally cloned cDNA library in the  
pSPORT1 vector (Life Technologies) was constructed at  
BioServe Biotechnology (Laurel MD) essentially following  
the protocols of the SuperScript plasmid System full  
details of which are contained in the manufacturer's  
instruction manual (http://www.lifetech.com/). First  
strand synthesis was carried out using a Not I  
primer-adaptor  
[5'-pGATAGTTCTAGATCGCGGGCGCCG(T)15-3']. EST analysis  
was performed on the unamplified library at the NIH  
Intramural Sequencing Center (NISC)."

## ORIGIN

Query Match 17.6%; Score 409; DB 5; Length 570;  
Best Local Similarity 100.0%; Pred. No. 1e-207;  
Matches 409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1628 GAAACAAGAGGAGATATAAGTGTATGAGTTGGGCAAGAGAGGACATTTGCATCTGT 1687  
DB 162 GAAACAAGAGGAGATATAAGTGTATGAGTTGGGCAAGAGAGGACATTTGCATCTGT 221  
QY 1688 CTGGCAACCGTCTACTCTACAATGTGAATGGACTGAAAGACAGAGCTTTGGTGGACCCCAA 1747  
DB 222 CTGGCAACCGTCTACTCTACAATGTGAATGGACTGAAAGACAGAGCTTTGGTGGACCCCAA 281  
QY 1748 AGAAACGGGCTACACAGAGTTAGTAAAGCTTTTGAAGAAAGCTGGAAGTAAATCCGAGA 1807  
DB 282 AGAAACGGGCTACACAGAGTTAGTAAAGCTTTTGAAGAAAGCTGGAAGTAAATCCGAGA 341  
QY 1808 TGRAGGCTTGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1867  
DB 342 TGAAGGCTTGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 401  
QY 1868 GAGGAGAAAGGGGAGCTGGAGCCAGTTTCAAGCTGTGGCAGCAAGAGAGAGAAATGAAAA 1927  
DB 402 GAGGAGAAAGGGGAGCTGGAGCCAGTTTCAAGCTGTGGCAGCAAGAGAGAGAAATGAAAA 461  
QY 1928 TGCCTGCAAGAGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCTCCGAGACACAGG 1987  
DB 462 TGCCTGCAAGAGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCTCCGAGACACAGG 521  
QY 1988 ATGCAGAGAGAGAGAGATCAATATTCATCATGACCCCGGAGCTCAC 2036  
DB 522 ATGCAGAGAGAGAGAGATCAATATTCATCATGACCCCGGAGCTCAC 570

## RESULT 46

BQ639558  
LOCUS  
DEFINITION  
17000600180518 GRN\_PRENU Homo sapiens cDNA 5', mRNA sequence.  
BQ639558  
VERSION  
BQ639558.1 GI:47368755  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1. (bases 1 to 740)  
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,

Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R., Lebkowski, J and Stanton, L.W.	
Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation	
Nat. Biotechnol. 22 (6), 707-716 (2004)	
Contact: Brandenberger R Regenerative Medicine Geron Corporation 230 Constitution Drive, Menlo Park, CA 94025, USA	
Tel: 650 473 8658	
Fax: 650 473 7760	
Email: tbrandenberger@geron.com	
Insert Length: 740	Std Error: 0.00.

```

FEATURES
source
Location/Qualifiers
1. .740
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cell, retinoic acid and
mitogen-treated hES cell line H7"
/clone_lib="GRN_PRENEU"
/note="oligo dt primed, full-length enriched cDNA library
from hES cell line H7 (p29) maintained in feeder-free
conditions. Embryoid bodies were generated in the presence
of all-trans retinoic acid and mitogens."
```

ORIGIN	Query Match	17.3%;	Score 402;	DB 7;	Length 740;
	Best Local Similarity	99.8%;	Pred. No. 6.4e-204;		
	Matches 452;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	49	GCACGACGACGGCTCCGGCAGCGGTAGCACAGTGGCGGCGAGCAGAGCCCGGGGCC	108		
Db	1	GCACGACGACGGCTCCGGCAGCGGTAGCACAGTGGCGGCGAGCAGAGCCCGGGGCC	60		
Qy	109	GGAGAGAGCAAAGCATGGAGGACACAAAGATGGAGGAAGGCGGACTCTCGGGAATT	168		
Db	61	GGAGAGAGCAAAGCATGGAGGACACAAAGATGGAGGAAGGCGGACTCTCAGGAATT	120		
Qy	169	CATTCTTCACGTGGTTTATGTGATTGCAATGTCTGGCGCTCTGGACATCTGTAGCTGCG	228		
Db	121	CATTCTTCACGTGGTTTATGTGATTGCAATGTCTGGCGCTCTGGACATCTGTAGCTGCG	180		
Qy	229	TTTGGTTTGATCTTGTTGACTATGAGGAAGTCTTAGGAAACTAGGAATCTATGATGCTG	288		
Db	181	TTTGGTTTGATCTTGTTGACTATGAGGAAGTCTTAGGAAACTAGGAATCTATGATGCTG	240		
Qy	289	ATGGTGATGGAGATTTTCATGTGATGATGTCACAAAGTTTTATTAGGACTTAAAGAGAGAT	348		
Db	241	ATGGTGATGGAGATTTTCATGTGATGATGTCACAAAGTTTTATTAGGACTTAAAGAGAGAT	300		
Qy	349	CTACTTCAGAGCCAGAGTCCGCCAGAGAGGCTGAGGCCACACTGAGCCCGAGGAGC	408		
Db	301	CTACTTCAGAGCCAGAGTCCGCCAGAGAGGCTGAGGCCACACTGAGCCCGAGGAGC	360		
Qy	409	AGGTTTCCTGTGGAGGAGAAACCCAGATATATCGAAGATGAAGCAAAAGAACAAATTCAGT	468		
Db	361	AGGTTTCCTGTGGAGGAGAAACCCAGATATATCGAAGATGAAGCAAAAGAACAAATTCAGT	420		
Qy	469	CCCTTCTCATAAATGGTACACGAGAACATG	501		
Db	421	CCCTTCTCATAAATGGTACACGAGAACATG	453		

[illegible]

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 629)  
 Tsai, J.Y. and Wistow, G.  
 Expressed sequence tag analysis of cultured primary human ocular  
 pericytes  
 Unpublished (2004)  
 Contact: Wistow G  
 Section on Molecular Structure and Function  
 National Eye Institute  
 6/331, NIH, Bethesda, MD 20892-2740, USA  
 Tel: 301 402 3452  
 Fax: 301 496 0078  
 Email: graeme@helix.nih.gov  
 Plate: 36 row: h column: 12  
 Seq primer: M13RPI reverse primer (ABI).  
 Location/Qualifiers  
 1. .629  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="hw36h12"  
 /cell\_type="pericytes"  
 /dev\_stage="Adult"  
 /lab\_host="EMDH10B"  
 /clone\_lib="Human primary human ocular pericytes."  
 /normalized="hw1"

Unamplified (nW) "note="Organ: Eye; Vector: pSport1; RNA was extracted from primary human pericytes in culture. A directionally cloned cDNA library in the pSport1 vector (Invitrogen) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's instruction manual (<http://www.lifetech.com/>). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGCACTAGTCTACATCGGCGGCGGCC(T)15-3']. cDNA was cloned in Not I/Sal I sites. EST analysis was performed at the National Sequencing Center (NISC) "

## ORIGIN

Query Match	17.1%;	Score 397;	DB 7;	Length 629;
Best Local Similarity	99.8%;	Pred. No. 3.1e-201;		
Matches 44%;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	985	AAGTTAAGAAAAGAAGCCTFAAACCTTTTAAATAAAATTTGATAAGACTATTAAAGCTGAAC	1044	
Db	182	AAGTTAAGAAAAGAAGCCTFAAACCTTTTAAATAAAATTTGATAAGACTATTAAAGCTGAAC	241	
Qy	1045	TTGATGCTGCAGAAAAAATCCCGTAAAGGGGAAAAAATTGAGGAAGCAGTGAATGCATTTA	1104	
Db	242	TTGATGCTGCAGAAAAAATCCCGTAAAGGGGAAAAAATTGAGGAAGCAGTGAATGCATTTA	301	
Qy	1105	AAGAACTAGTACGCAAAATACCCCTCAGAGTCCACGAGCAAGATATGGGAAGGCGCAGTGTG	1164	
Db	302	AAGAACTAGTACGCAAAATACCCCTCAGAGTCCACGAGCAAGATATGGGAAGGCGCAGTGTG	361	
Qy	1165	AGGATGATTTGGCTGCAGAAAGGAGGAAGTAAATGAGGTGCTACGTGGAGCCATCGAGACCT	1224	
Db	362	AGGATGATTTGGCTGCAGAAAGGAGGAAGTAAATGAGGTGCTACGTGGAGCCATCGAGACCT	421	
Qy	1225	ACCAAGAGGTGGCCAGCCTACCTGATGTCCTGCAGACCTGCTGAAGCTCAGTTTGAAGC	1284	
Db	422	ACCAAGAGGTGGCCAGCCTACCTGATGTCCTGCAGACCTGCTGAAGCTCAGTTTGAAGC	481	
Qy	1285	GTCGCTCAGACAGGCAACAATTTCTAGGTCATATGAGAGGTTCCCTGCTTACCTGCAGA	1344	
Db	482	GTCGCTCAGACAGGCAACAATTTCTAGGTCATATGAGAGGTTCCCTGCTTACCTGCAGA	541	
Qy	1345	GATTAGTTCAACTATTTCCCAATGATATCTTCCTTAAAAAATGACCTGGCGTGGATACC	1404	
Db	542	GATTAGTTCAACTATTTCCCAATGATATCTTCCTTAAAAAATGACCTGGCGTGGATACC	601	



```

/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC_72"
/notes="Organ: skin; Vector: pCMV-SF1"
Site_2: Sali; Cloned unidirectional
Average insert size 2 kb. Library
Technologies."

```

```
Query Match      17.0%; Score 395; DB 5; Length 503;
Best Local Similarity 100.0%; Pred. No. 3.5e-200;
Matches 395; Conservative 0; Mismatches 0; Indels
```

ORIGIN	Query Match	16.9%;	Score 392;	DB 5;	Length 1031;
	Best Local Similarity	100.0%;	Pred. No. 1.6e-198;		
	Matches 392;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	555	GAGGATGATCAGATTCTTATGGCGACTGATGTAGATGATAGATTGAGACCCCTGGAACCT	614		
Db	11	GAGGATGATCAGATTCTTATGGCGACTGATGTAGATGATAGATTGAGACCCCTGGAACCT	70		
QY	615	GAAGTATCTCATGAAGAAACCGAGCATAGATTACACGTGGAAGACAGATTTCACAGAC	674		
Db	71	GAAGTATCTCATGAAGAAACCGAGCATAGATTACACGTGGAAGACAGATTTCACAGAC	130		
QY	675	TGTAAATCAGGATATGGAAGAGATGATCTCTGAGCAGGAAAATCCAGATTTCCAGTGAACCA	734		
Db	131	TGTAAATCAGGATATGGAAGAGATGATCTCTGAGCAGGAAAATCCAGATTTCCAGTGAACCA	190		
QY	735	GTAGTAGAAGATGAAGAGATTGCACCATGATACAGATGATGAAATACCAAGTCTATGAG	794		
Db	191	GTAGTAGAAGATGAAGAGATTGCACCATGATACAGATGATGAAATACCAAGTCTATGAG	250		
QY	795	GAAACAAGCATATGAACCTCTAGAAAATGAAGGGATAGAAAATCACAGAAGTAAGTCTCT	854		
Db	251	GAAACAAGCATATGAACCTCTAGAAAATGAAGGGATAGAAAATCACAGAAGTAAGTCTCT	310		
QY	855	CCCCCTCAGGATTAATCTGTAGAAAGATTCCACAGTAAATTGTAGAAGAAAGTAAGCATTTTT	914		
Db	311	CCCCCTCAGGATTAATCTGTAGAAGATTCCACAGTAAATTGTAGAAGAAAGTAAGCATTTTT	370		
QY	915	CCTGTGAAGAACACGAGGAAGTACCACCCAGA	946		
Db	371	CCTGTGAAGAACACGAGGAAGTACCACCCAGA	402		

RESULT 51	EST 14-JUN-2001
BI035330	linear
LOCUS	553 bp mRNA
Accession:	BI035330 Homo sapiens cDNA.
Version:	0.224001.624=03.NR2003
Definition:	EST
Accession:	U00000.1 Homo sapiens mRNA sequence.

RECEIVED  
VERSION  
BI035330.1  
GI:14441956

SOURCE Homo sapiens (human)

T  
WTH MOC htrp. /mac nci nih gov/  
(bases I 031

**NATIONAL INDIANCES OF ADOLESCENT  
PUBLISHED (1999)**

Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Agencourt Bioscience Corporation

found through the I.M.A.G.E. Consortium/LLNL at:

Plate: LLAM13340 ROW: 1 COLUMN: 08

STATTPNØ/NOTPBOI  
1021

```
/organelle= homo sapiens
/mot time="mrna"
```

```
/w_ajax_content.php  
/clone="IMAGE:6065047"
```

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV2&t2=QV2-NN2003-230401-624-a03&t3=2001-04-23&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence stop: 525.  
 Location/Qualifiers  
 1..553  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="NN2003"  
 /note="Organ: nervous normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

#### FEATURES

source

#### ORIGIN

Query Match 16.7%; Score 389; DB 4; Length 553;  
 Best Local Similarity 99.6%; Pred. No. 66-197;  
 Matches 489; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 332 AGGACTTAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACA 391  
 DB 17 AGGACTTAAGAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACA 76

QY 392 CACTGAGCCGAGGAGCAGGTTCTCTGTGGAGGAGAACCCAGAGATATCGAAGTGAAGC 451  
 DB 77 CACTGAGCCGAGGAGCAGGTTCTCTGTGGAGGAGAACCCAGAGATATCGAAGTGAAGC 136

QY 452 AAAAGAACAAATTCAGTCCCTCTCCATGATGTGTACGCGAGACATGTTGAGGGAGA 511  
 DB 137 AAAAGAACAAATTCAGTCCCTCTCCATGATGTGTACGCGAGACATGTTGAGGGAGA 196

QY 512 AGACTTGCACACAGAGATGAGCCACAGGAGAACCCACACAGAGAGATGATGTTCT 571  
 DB 197 AGACTTGCACACAGAGATGAGCCACAGGAGAACCCACACAGAGAGATGATGTTCT 256

QY 572 TATGGCAGTGTATGATGATGATTTGAGACCTCGAACCTGAAAGTATCTCATGAAGA 631  
 DB 257 TATGGCAGTGTATGATGATGATTTGAGACCTCGAACCTGAAAGTATCTCATGAAGA 316

QY 632 AACCGAGCATAGTTACACGTTGGAAGAGACAGTTTCAACAGACTGTAATCAGATATGGA 691  
 DB 317 AACCGAGCATAGTTACACGTTGGAAGAGACAGTTTCAACAGACTGTAATCAGATATGGA 376

QY 692 AGAGATGATCTGTGACAGAGAAATCCAGATTCAGTGAAACAGTGTAGAGAGATGAAG 751  
 DB 377 AGAGATGATCTGTGACAGAGAAATCCAGATTCAGTGAAACAGTGTAGAGAGATGAAG 436

QY 752 ATTGCCACATGATACAGATGATGTAAACATACCAAGTCTATGAGGAACAAGCAGTATATGA 811  
 DB 437 ATTGCCACATGATACAGATGATGTAAACATACCAAGTCTATGAGGAACAAGCAGTATATGA 496

QY 812 ACCTCTAGAAA 822  
 DB 497 ACCTCTAGAAA 507

#### RESULT 52

BE084010/c  
 LOCUS  
 DEFINITION PMO-BT0651-150300-002-905 BT0651 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BE084010  
 VERSION BE084010.1 GI:8474363  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

#### REFERENCE

##### AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 409)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

##### TITLE

Shotgun sequencing of the human transcriptome with ORF expressed

##### JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

##### MEDLINE

20202663

##### PUBMED

10737800

##### COMMENT

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil.  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PMO-BT0651-150300-002-g05&t3=2000-03-15&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 18

High quality sequence stop: 409.

#### FEATURES

source

1..409  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="BT0651"

/note="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

#### ORIGIN

Query Match 16.6%; Score 385; DB 2; Length 409;  
 Best Local Similarity 100.0%; Pred. No. 8.1e-195;  
 Matches 385; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 347 ATCTACTTCAGAGCCAGAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGA 406  
 DB 402 ATCTACTTCAGAGCCAGAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGA 343

QY 407 GCAGGTTCTGTGGAGGAGAACCCAGAGATATCGAAGATGAAGCAAAAGCAAAATTC 466  
 DB 342 GCAGGTTCTGTGGAGGAGAACCCAGAGATATCGAAGATGAAGCAAAAGCAAAATTC 283

QY 467 GTCCCTTCTCCATGAATGTTACAGCGAGAACCATGTTGAGGAGAGAGCTTCACACAGA 526  
 DB 282 GTCCCTTCTCCATGAATGTTACAGCGAGAACCATGTTGAGGAGAGAGCTTCACACAGA 223

QY 527 AGATGGACCCAGGAGAACCAACAGAGGATGATGATTTCTTATGGCGACTGATGT 586  
 DB 222 AGATGGACCCAGGAGAACCAACAGAGGATGATGATTTCTTATGGCGACTGATGT 163

QY 587 AGATGATAGATTGAGACCTTGGAACTCTGAAGTATCTCATGAAGAAACCGAGCATAGTTA 646  
 DB 162 AGATGATAGATTGAGACCTTGGAACTCTGAAGTATCTCATGAAGAAACCGAGCATAGTTA 103

QY 647 CCACGTGGAAGAGACAGTTTCAAGAGCTCTAATCAGGATATGGAAGAGATGATGTTCTGA 706  
 DB 102 CCACGTGGAAGAGACAGTTTCAAGAGCTCTAATCAGGATATGGAAGAGATGATGTTCTGA 43

QY 707 GCAGGAAATTCAGGATTCAGTGA 731

```
Db 42 GCAGAAATCCAGATTCAGTGA 18
|||||
BU838423 930 bp mRNA linear EST 16-OCT-2002
AGENCOURT 8123161 Lupski dorsal_root_ganglion Homo sapiens cDNA
clone IMAGE:6177038 5', mRNA sequence.
BU838423
VERSION BU838423.1 GI:24022818
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLML)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLML at:
http://image.llnl.gov
Plate: LLML3554 row: f column: 15
High quality sequence start: 66
High quality sequence stop: 549.
Location/Qualifiers
1..930
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6177038"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_lib="Lupski_dorsal_root_ganglion"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGGCTCG-3' and
5'-GACTAGTCTAGATCGGAGCGGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.7 Kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
ORIGIN
Query Match 16.5%; Score 383; DB 5; Length 930;
Best Local Similarity 99.4%; Pred. No. 1.1e-193;
Matches 533; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CGGACCGTGCATGCGCCAGCGTAGATCCAGAGCAGCGGCGCACAGCAGCAGCG 60
|||||
Db 138 CGGACCGTGCATGCGCCAGCGTAGAATGCCAAGAGCAGCGGCGCACAGCAGCAGCG 197
|||||
QY 61 GCTCCGGCAGCGGTAGCAGTGGCGGAGCAGCAGCGCCCGGCGGAGAGACAA 120
|||||
Db 198 GCTCCGGCAGCGGTAGCAGTGGCGGAGCAGCAGCGCCCGGCGGAGAGACAA 257
|||||
QY 121 AGCATGGAGGACACAAGAAATGGGAGAAAGCGCGACTCTCGGGAACCTTCATCTTCACGT 180
|||||
Db 258 AGCATGGAGGACACAAGAAATGGGAGAAAGCGCGACTCTCAGGAACCTTCATCTTCACGT 317
|||||
QY 181 GGTATTGTTGATTCGATTCGTCGGGCTGTGACATCTGTAGCTGTGCTGTTGATC 240
|||||
```

```
Db 318 GGTATTGTTGATTCGATTCGTCGGGCTGTGACATCTGTAGCTGTGCTGTTGATC 377
|||||
QY 241 TTGTTGACTATCAGGAAGTTCTAGGAAACTAGGAATCTATGATCTGATGTTGAGGAG 300
|||||
Db 378 TTGTTGACTATCAGGAAGTTCTAGGAAACTAGGAATCTATGATCTGATGTTGAGGAG 437
|||||
QY 301 ATTTTGTGATGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 360
|||||
Db 438 ATTTTGTGATGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 497
|||||
QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGTTCTTGTGG 420
|||||
Db 498 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGGTTCTTGTGG 557
|||||
QY 421 AGCGAAGACCCAGAGATATCGAAGATGAAGCAAGCAAAATTCAGTCCCTTCTCCATG 480
|||||
Db 558 AGCGAAGACCCAGAGATATCGAAGATGAAGCAAGCAAAATTCAGTCCCTTCTCCATG 617
|||||
QY 481 AAATGTTACACGCGAGAACATGTTGAGGGAGAGACTTGCACCAAGAGATGGACCC 536
|||||
Db 618 AAATGTTACACGCGAGAACATGTTGAGGGAGAGACTTGCACCAAGAGATGGACCC 673
|||||
RESULT 54
CN368831 468 bp mRNA linear EST 16-MAY-2004
LOCUS 17000600846567 GRN_PRENU Homo sapiens cDNA 5', mRNA sequence.
DEFINITION CN368831
ACCESSION CN368831.1 GI:47368765
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 468)
JOURNAL Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M. S., Mandalam, R.,
COMMENT Lebkowski, J. and Stanton, L. W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 468 Std Error: 0.00.
Location/Qualifiers
1..468
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cell, retinoic acid and
mitogen-treated hES cell line H7"
/clone_lib="GRN_PRENU"
/notes="oligo dT primed, full-length enriched cDNA library
from hES cell line H7 (p29) maintained in feeder-free
conditions. Embryoid bodies were generated in the presence
of all-trans retinoic acid and mitogens."
ORIGIN
Query Match 16.4%; Score 382; DB 7; Length 468;
Best Local Similarity 100.0%; Pred. No. 3.4e-193;
Matches 382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 251 TGAGGAAGTCTAGGAAACTAGGAATCTATGATGCTGATGATGAGATTGATGT 310
|||||
Db 7 TGAGGAAGTCTAGGAAACTAGGAATCTATGATGCTGATGATGAGATTGATGT 66
|||||
QY 311 GGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGCAGTCCC 370
|||||
```



Db 67 GGATGATGCCAAGTTTATTATAGGACTTAAAGAGAGATCTACTTCAGAGCCAGGTCCTCC 126  
 QY 371 GCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCCTGTGGAGGCGAGAAC 430  
 Db 127 GCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCCTGTGGAGGCGAGAAC 186  
 QY 431 CCAGATATTCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTCTCCATGAATGCTACA 490  
 Db 187 CCAGATATTCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTCTCCATGAATGCTACA 246  
 QY 491 CGCAGAACATGTTGGGAGAGAGACTTGCACAGAGATGAGCCACAGAGAACCCACA 550  
 Db 247 CGCAGAACATGTTGGGAGAGAGACTTGCACAGAGATGAGCCACAGAGAACCCACA 306  
 QY 551 ACAGAGGATGATGAGTTTCTTATGGCGACTGATGTAGATGATAGATTGAGACCCCTGGA 610  
 Db 307 ACAAGAGGATGATGAGTTTCTTATGGCGACTGATGTAGATGATAGATTGAGACCCCTGGA 366  
 QY 611 ACCTGAAGTATCTCATGAAGAA 632  
 Db 367 ACCTGAAGTATCTCATGAAGAA 388

RESULT 55  
 LOCUS CA311352 720 bp mRNA linear EST 04-NOV-2002  
 DEFINITION UI-CF-FNO-aff-d-19-0-UI.81 UI-CF-FNO Homo sapiens cDNA clone  
 ACCESSION CA311352  
 VERSION CA311352.1 GI:24529450  
 KEYWORDS EST.

ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 720)  
 Bonaldo, M.F., Lennon, G. and Soares, M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)  
 97044477  
 889546  
 Contact: McCray, PB  
 McCray Lab  
 University of Iowa  
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
 Tel: 319 356 4866  
 Fax: 319 356 7171  
 Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com) or from Open Biosystems  
 (www.openbiosystems.com).  
 Seq primer: M13 FORWARD  
 POLYA=Yes.

Location/Qualifiers  
 1. .720  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-CF-FNO-aff-d-19-0-UI"  
 /tissue\_type="Human Lung Epithelial cells"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-CF-FNO"  
 /notes="Organ: Lung; Vector: p7T3-Pac (Pharmacia) with a  
 modified polylinker; Site: 1: EcoR I; Site 2: Not I;  
 UI-CF-FNO is a subtracted cDNA library derived from two  
 normalized Human lung epithelial cell libraries (EN1 and  
 DU1) The library was subtracted according to according to  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. For additional information, contact:  
 bent-soares@uiowa.edu  
 TAG\_L18UB=lung Epithelial Cells Tissue nos 359-368  
 TAG\_L18=UI-CF-FNO  
 TAG\_SEQ=GGCTGTAGGC

## ORIGIN

Query Match 16.3%; Score 379; DB 6; Length 720;  
 Best Local Similarity 99.8%; Pred.No. 1.5e-191;  
 Matches 429; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1895 CACGCTGTGGCAGCAAGGAAGAAATGAAATSCCTGCAAGAGACTCTCTAAACCTG 1954  
 Db 51 CACGCTGTGGCAGCAAGGAAGAAATGAAATSCCTGCAAGAGACTCTCTAAACCTG 110  
 QY 1955 TACCTTACTAGAAAGTTCCCGAGACACAGGATGAGAGAGGACAGATCAATATTC 2014  
 Db 111 TACCTTACTAGAAAGTTCCCGAGACACAGGATGAGAGAGGACAGATCAATATTC 170  
 QY 2015 CATATGCAACCCGCGGACTCACGTGTGGCCGACACAGGGCCCAAACTGCAGGCTCG 2074  
 Db 171 CATATGCAACCCGCGGACTCACGTGTGGCCGACACAGGGCCCAAACTGCAGGCTCG 230  
 QY 2075 AATGACCTGGGCTTGGTGATTCCTCAAGGAAGGCTGCAAGATTGATGTCCTCAACGAGAC 2134  
 Db 231 AATGACCTGGGCTTGGTGATTCCTCAAGGAAGGCTGCAAGATTGATGTCCTCAACGAGAC 290  
 QY 2135 CAGGACCTGGGAGGAAGGCAAGGTGCTCATCTTTGATGACTCTTTGAGCAGGAGTATG 2194  
 Db 291 CAGGACCTGGGAGGAAGGCAAGGTGCTCATCTTTGATGACTCTTTGAGCAGGAGTATG 350  
 QY 2195 GCAGGATGCTCATCTTTCCGGCTGATATTCATCGTGGATGTGGCATCCGGAACCTGAC 2254  
 Db 351 GCAGGATGCTCATCTTTCCGGCTGATATTCATCGTGGATGTGGCATCCGGAACCTGAC 410  
 QY 2255 ACCACACAGAGACGAGCCTTCCAGCAATTTAGCATGAATTCATCAAGCTTGGGAAC 2314  
 Db 411 ACCACACAGAGACGAGCCTTCCAGCAATTTAGCATGAATTCATCAAGCTTGGGAAC 470  
 QY 2315 TCTGGAGAGA 2324  
 Db 471 TCTGGAGAGA 480

RESULT 56  
 LOCUS CN368826 626 bp mRNA linear EST 16-MAY-2004  
 DEFINITION 17000423444515 GRN\_EB Homo sapiens cDNA 5', mRNA sequence.  
 ACCESSION CN368826  
 VERSION CN368826.1 GI:47368760  
 KEYWORDS EST.

ORGANISM Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 626)  
 Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,  
 Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,  
 Lebkowski, J. and Stanton, L.W.  
 Transcriptome characterization elucidates signaling networks that  
 control human ES cell growth and differentiation  
 Nat. Biotechnol. 22 (6), 707-716 (2004)  
 Contact: Brandenberger R  
 Regenerative Medicine  
 Genon Corporation  
 230 Constitution Drive, Menlo Park, CA 94025, USA  
 Tel: 650 473 8658  
 Fax: 650 473 7760  
 Email: rbrandenberger@genon.com  
 Insert Length: 626 Std Error: 0.00.  
 Location/Qualifiers  
 1. 626  
 /organism="Homo sapiens"

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## FEATURES

## source

mol\_type="mRNA"  
db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cells, embryoid bodies  
derived from H1, H7 and H9 cells"  
/clone\_lib="GRN\_EB"  
/note="oligo dt primed, full-length enriched cDNA library  
from embryoid body outgrowths derived from hES cell lines  
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free  
conditions."

ORIGIN  
Query Match 15.9%; Score 370; DB 7; Length 626;  
Best Local Similarity 99.8%; Pred. No. 1e-186;  
Matches 420; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CGGACCGTGCATGGCCAGCGTAAGATGCCAAGACGAGCGGCAACAGCAGCAGCAGCG 60  
Db CGGACCGTGCATGGCCAGCGTAAGATGCCAAGACGAGCGGCAACAGCAGCAGCAGCG 265  
Qy 61 GCTCCGCGAGGTAGCAGCAGTGGCGGAGCGAGCAGCGCGCGCGCGCGCGCGAGACAA 120  
Db GCTCCGCGAGGTAGCAGCAGTGGCGGAGCGAGCAGCGCGCGCGCGCGCGAGACAA 325  
Qy 121 AGCATGGAGGACACAAAGATGGGAGGAAAGCGGACTCTCGGGAACCTTCATTTTCAGT 180  
Db AGCATGGAGGACACAAAGATGGGAGGAAAGCGGACTCTCAGGAACTTCATTTTCAGT 385  
Qy 181 GGTATTGATGATTCGATTCGCGGCGTCTGGACATCTGTAGCTGTCTGTTGTTGATC 240  
Db GGTATTGATGATTCGATTCGCGGCGTCTGGACATCTGTAGCTGTCTGTTGTTGATC 445  
Qy 241 TTGTTGACTATGAGGAGTCTAGGAAACTAGGAATCTATGATGCTGATGCTGATGGAG 300  
Db TTGTTGACTATGAGGAGTCTAGGAAACTAGGAATCTATGATGCTGATGCTGATGGAG 505  
Qy 301 ATTTTGTGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 360  
Db ATTTTGTGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 565  
Qy 361 CAGCAGTCCCGCAGAGGCTGAGCAGACACACTGAGCCGCGAGGAGCAGGTCCTGTGG 420  
Db CAGCAGTCCCGCAGAGGCTGAGCAGACACACTGAGCCGCGAGGAGCAGGTCCTGTGG 625  
Qy 421 A 421  
Db 626 A 626

RESULT 57  
BE876059  
LOCUS BE876059 1048 bp mRNA linear EST 20-OCT-2000  
DEFINITION 601485742F1 NIH\_MGC\_69 Homo sapiens cDNA clone IMAGE:3887962 5',  
mRNA sequence.  
ACCESSION BE876059  
VERSION BE876059.1 GI:10324835  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: DCTD/DTF/Gazdar  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov

Plate: LLAM9667 row: d column: 11  
High quality sequence stop: 670.  
Location/Qualifiers  
1. 1048  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3887962"  
/tissue\_type="large cell carcinoma, undifferentiated"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 69"  
/note="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.1 Kb. Library constructed by Life  
Technologies."

ORIGIN  
Query Match 14.2%; Score 329; DB 2; Length 1048;  
Best Local Similarity 99.5%; Pred. No. 1.3e-164;  
Matches 429; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 CGGACCGTGCATGGCCAGCGTAAGATGCCAAGACGAGCGGCAACAGCAGCAGCAGCG 60  
Db CGGACCGTGCATGGCCAGCGTAAGATGCCAAGACGAGCGGCAACAGCAGCAGCAGCG 228  
Qy 61 GCTCCGCGAGGTAGCAGCAGTGGCGGAGCGAGCAGCGCGCGCGCGCGAGACAA 120  
Db GCTCCGCGAGGTAGCAGCAGTGGCGGAGCGAGCAGCGCGCGCGCGCGAGACAA 288  
Qy 121 AGCATGGAGGACACAAAGATGGGAGGAAAGCGGACTCTCGGGAACCTTCATTTTCAGT 180  
Db AGCATGGAGGACACAAAGATGGGAGGAAAGCGGACTCTCAGGAACTTCATTTTCAGT 348  
Qy 181 GGTATTGATGATTCGATTCGCGGCGTCTGGACATCTGTAGCTGTCTGTTGTTGATC 240  
Db GGTATTGATGATTCGATTCGCGGCGTCTGGACATCTGTAGCTGTCTGTTGTTGATC 408  
Qy 241 TTGTTGACTATGAGGAGTCTAGGAAACTAGGAATCTATGATGCTGATGCTGATGGAG 300  
Db TTGTTGACTATGAGGAGTCTAGGAAACTAGGAATCTATGATGCTGATGCTGATGGAG 468  
Qy 301 ATTTTGTGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 360  
Db ATTTTGTGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 528  
Qy 361 CAGCAGTCCCGCAGAGGCTGAGCAGACACACTGAGCCGCGAGGAGCAGGTCCTGTGG 420  
Db CAGCAGTCCCGCAGAGGCTGAGCAGACACACTGAGCCGCGAGGAGCAGGTCCTGTGG 588  
Qy 421 AGGAGAGACCC 431  
Db 589 AGGAGAGACCC 599

RESULT 58  
AI267311/c  
LOCUS AI267311 574 bp mRNA linear EST 17-NOV-1998  
DEFINITION aq63e03.x1 Stanley Frontal SN pool 2 Homo sapiens cDNA clone  
IMAGE:2035612 similar to SW:ASPH HUMAN Q12797 ASPATYL/ASPARAGINYL  
BETA-HYDROXYLASE ; mRNA sequence.  
ACCESSION AI267311  
VERSION AI267311.1 GI:3886478  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS 1 (bases 1 to 574)  
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,  
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,  
Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F.,  
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.  
WashU-NCI human EST Project  
TITLE

JOURNAL  
COMMENT

Unpublished (1997)  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: estewartson.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Seq primer: -400P from Gibco  
 High quality sequence stop: 439.

FEATURES  
source

Location/Qualifiers  
 1..574  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2035612"  
 /tissue types="frontal lobe (see description)"  
 /clone host="Stanley Frontal SN pool 2"  
 /note="Organ: brain; Vector: pOR2.1-TOPO (Invitrogen);  
 Site 1: EcoRI; Total RNA (purified with Trizol and DNaseI  
 before use) was reverse transcribed using a modified  
 oligo-dr primer containing RsaI and HindIII sites.  
 Double-stranded cDNA was digested with RsaI, resulting in  
 blunt ended cDNA of an average 0.1-2 kb in length.  
 Digested cDNA was split into two sets, one used as is as  
 the driver, the other set was split in half again and each  
 half linked to a different adaptor  
 (5'-TCGAGCGGCGCGCGGAGGT-3' or 5'-  
 AGGCGGTGTCGCGGCGGT-3'), to be used as tester.  
 Subtraction was performed using the Clontech PCR Select  
 cDNA subtraction kit. Pool of two schizophrenics, male age  
 44 and female age 56 (S-116, S-118) subtracted by pool of  
 two mentally normal male individuals ages 41 and 53  
 (S-124, S-141). Tissues were obtained from the Stanley  
 Neuropathology Consortium (www.stanleylab.org). Library  
 constructed and subtracted by Dr. Nancy Johnston [(410)  
 614-3918, nlj@welchlink.weich.jhu.edu]."

## ORIGIN

Query Match 14.1%; Score 328; DB 1; Length 574;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-164;  
 Matches 328; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 GGAATTCATCTTCACGTGTTATGTTGATTGCTGCGGCGCTCGGACATCTGTA 221  
 DB 345 GGAATTCATCTTCACGTGTTATGTTGATTGCTGCGGCGCTCGGACATCTGTA 286  
 QY 222 GCTGTCGTTGGTTGATCTTGTGACTATGAGGAGTTCTAGAAAACCTAGGAATCTAT 281  
 DB 285 GCTGTCGTTGGTTGATCTTGTGACTATGAGGAGTTCTAGAAAACCTAGGAATCTAT 226  
 QY 282 GATGCTGATGTTGATGAGATTTTGTGATGATGATGCCAAAGTTTATTAGGACTTAAA 341  
 DB 225 GATGCTGATGTTGATGAGATTTTGTGATGATGATGCCAAAGTTTATTAGGACTTAAA 166  
 QY 342 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCC 401  
 DB 165 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCC 106  
 QY 402 GAGGAGCAGGTTCTCTGGAGCAGAACCCAGAGATATCGAATGAAGCAAAAGACAA 461  
 DB 105 GAGGAGCAGGTTCTCTGGAGCAGAACCCAGAGATATCGAATGAAGCAAAAGACAA 46  
 QY 462 ATTCAGTCCCTTCTCCATGAATGTAC 489  
 DB 45 ATTCAGTCCCTTCTCCATGAATGTAC 18

RESULT 59  
 CN368818  
 LOCUS

CN368818 393 bp mRNA linear EST 16-MAY-2004

DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

17000600014160 GRN\_PRENU Homo sapiens cDNA 5', mRNA sequence.  
 CN368818  
 CN368818.1 GI:47368752  
 EST.  
 Homo sapiens (human)

REFERENCE  
 AUTHORS

1 (bases 1 to 393)  
 Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,  
 Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M.S., Mandalam, R.,  
 Lebkowski, J. and Stanton, L.W.  
 Transcriptome characterization elucidates signaling networks that  
 control human ES cell growth and differentiation  
 Nat. Biotechnol. 22 (6), 707-716 (2004)  
 Contact: Brandenberger R  
 Regenerative Medicine  
 Geron Corporation  
 230 Constitution Drive, Menlo Park, CA 94025, USA  
 Tel: 650 473 8658  
 Fax: 650 473 7760  
 Email: rbrandenberger@geron.com  
 Insert Length: 393 Std Error: 0.00.

## TITLE

JOURNAL  
 COMMENTFEATURES  
 source

Location/Qualifiers  
 1..393  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="embryonic stem cell, retinoic acid and  
 mitogen-treated HES cell line H7"  
 /clone\_lib="GRN\_PRENU"  
 /notes="Oligo dT primed, full-length enriched cDNA library  
 from HES cell line H7 (p29) maintained in feeder-free  
 conditions. Embryoid bodies were generated in the presence  
 of all-trans retinoic acid and mitogens."

## ORIGIN

Query Match 13.9%; Score 324; DB 7; Length 393;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-162;  
 Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 GGAATTCATCTTCACGTGTTATGTTGATTGCTGCGGCGCTCGGACATCTGTA 221  
 DB 70 GGAATTCATCTTCACGTGTTATGTTGATTGCTGCGGCGCTCGGACATCTGTA 129  
 QY 222 GCTGTCGTTGGTTGATCTTGTGACTATGAGGAGTTCTAGAAAACCTAGGAATCTAT 281  
 DB 130 GCTGTCGTTGGTTGATCTTGTGACTATGAGGAGTTCTAGAAAACCTAGGAATCTAT 189  
 QY 282 GATGCTGATGTTGATGAGATTTTGTGATGATGATGCCAAAGTTTATTAGGACTTAAA 341  
 DB 190 GATGCTGATGTTGATGAGATTTTGTGATGATGATGCCAAAGTTTATTAGGACTTAAA 249  
 QY 342 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCC 401  
 DB 250 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCC 309  
 QY 402 GAGGAGCAGGTTCTCTGGAGCAGAACCCAGAGATATCGAATGAAGCAAAAGACAA 461  
 DB 310 GAGGAGCAGGTTCTCTGGAGCAGAACCCAGAGATATCGAATGAAGCAAAAGACAA 369  
 QY 462 ATTCAGTCCCTTCTCCATGAATG 485  
 DB 370 ATTCAGTCCCTTCTCCATGAATG 393

RESULT 60  
 CN481268  
 LOCUS

DEFINITION  
 hw05d04.y1 Human primary human ocular pericytes. Unamplified (hw)  
 Homo sapiens cDNA clone hw05d04 5', mRNA sequence.  
 ACCESSION  
 CN481268  
 VERSION  
 CN481268.1 GI:46562772

CN481268 580 bp mRNA linear EST 26-APR-2004

KEYWORDS SOURCE ORGANISM  REFERENCE AUTHORS TITLE  JOURNAL COMMENT	EST. Homo sapiens (human)		Db	 515 CAGCAGTCCGCC 527		
	Homo sapiens					
	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
	1 (bases 1 to 580)					
	Tsai,J.Y. and Wistow,G.					
	Expressed sequence tag analysis of cultured primary human ocular pericytes					
	Unpublished (2004)					
	Contact: Wistow G					
	Section on Molecular Structure and Function National Eye Institute 6/331, NIH, Bethesda, MD 20892-2740, USA Tel: 301 402 3452 Fax: 301 496 0078 Email: graeme@helix.nih.gov Plate: 05 row: d column: 04 Seq primer: M13rpl reverse primer (ABI).					
	Location/Qualifiers					
FEATURES source	1..580 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="hw05d04" /cell_type="pericytes" /dev_stage="Adult" /lab_host="EMDH10B" /clone_lib="Human primary human ocular pericytes." Unamplified (hw) /note="Organ: Eye; Vector: pSport1; RNA was extracted from primary human pericytes in culture. A directionally cloned cDNA library in the pSPORT1 vector (Invitrogen) was constructed at Bioserve Biotechnology (Laurel MD) essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor [5'-pGACTAGTTCTAGTCGCGCGCCCT(r)15-3']. cDNA was cloned in Not I/Sal I sites. EST analysis was performed at the NIH Intramural Sequencing Center (NISC)."					
ORIGIN	Query Match 13.9%; Score 322; DB 7; Length 580; Best Local Similarity 99.7%; Pred.No.6.9e-161; Matches 372; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		QY			
	1 CGGACCGTGCATGGCCCAAGCGTAAGATGCCAAGAGCAGCGCGGCAACAGCAGCAGCGG 60					
	155 CGGACCGTGCATGGCCCAAGCGTAAGATGCCAAGAGCAGCGCGGCAACAGCAGCAGCGG 214					
	61 GTCGCGGACGGGTAGCAGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120					
	215 GTCGCGGACGGGTAGCAGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 274					
	121 AGCATGGAGGACACAGAATGGGAGGAAGCGGACTCTGGGAACCTTCATCTTCACGT 180					
	275 AGCATGGAGGACACAGAATGGGAGGAAGCGGACTCTCAGGAACCTTCATCTTCACGT 334					
ORIGIN	Query Match 13.8%; Score 321; DB 1; Length 527; Best Local Similarity 100.0%; Pred.No.2.3e-160; Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		QY			
	626 TGAAGAAACCGAGCATAGTTACCAACGTGGGAAGAGACAGTTTCAACAGACTGTATCAGGA 685					
	1 TGAAGAAACCGAGCATAGTTACCAACGTGGGAAGAGACAGTTTCAACAGACTGTATCAGGA 60					
	686 TATGGAAGAGATGATGTCGTGACGAGGAAATCCAGATTCCAGTGAACACGAGTAGTAGAGA 745					
	61 TATGGAAGAGATGATGTCGTGACGAGGAAATCCAGATTCCAGTGAACACGAGTAGTAGAGA 120					
	746 TGAAGATTGCACCATGATACAGATGATGTAACATACCAAGTCTATCAGGAAACAAGCAGT 805					
	121 TGAAGATTGCACCATGATACAGATGATGTAACATACCAAGTCTATCAGGAAACAAGCAGT 180					
ORIGIN	Query Match 13.8%; Score 321; DB 1; Length 527; Best Local Similarity 100.0%; Pred.No.2.3e-160; Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		QY			
	806 ATATGAACCTCTAGAAAATGAAGGATAGAAAATCACAAGAAGTAAGTCTCTCCCTCAGGA 865					
	181 ATATGAACCTCTAGAAAATGAAGGATAGAAAATCACAAGAAGTAAGTCTCTCCCTCAGGA 240					
	866 TAATCTCTAGAGAATTCACAGGTAATTTGTAGAAGAGTAAGCATTTTCTCTGTGGAAGA 925					
	241 TAATCTCTAGAGAATTCACAGGTAATTTGTAGAAGAGTAAGCATTTTCTCTGTGGAAGA 300					
	926 ACAGCAGGAAGTACCACCAGA 946					
	301 ACAGCAGGAAGTACCACCAGA 321					

```
RESULT 62
AY405268      2023 bp   DNA       linear      GSS 16-DEC-2003
LOCUS         Pan troglodytes ASPH gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION    genomic survey sequence.
ACCESSION     AY405268
VERSION       AY405268.1   GI:39761242
KEYWORDS      GSS.
SOURCE        Pan troglodytes (chimpanzee)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
AUTHORS       1 (bases 1 to 2023)
              Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
              Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,
              Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
              Adams, M.D. and Cargill, M.
TITLE         Inferring nonneutral evolution from human-chimp-mouse orthologous
              gene trios
JOURNAL       Science 302 (5652), 1960-1963 (2003)
PUBMED        14671302
REFERENCE     2 (bases 1 to 2023)
              Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
              Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,
              Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
              Adams, M.D. and Cargill, M.
TITLE         Direct Submission
JOURNAL       Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
              Rockville, MD 20850, USA
COMMENT       This sequence was made by sequencing genomic exons and ordering
              them based on alignment
FEATURES      Location/Qualifiers
              source          1..2023
                        /organism="Pan troglodytes"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9598"
                        <1..52023
                        /gene="ASPH"
                        /locus_tag="HCM2167"
              gene           1..52023
              ORIGIN
Query Match      13.6%; Score 316; DB 9; Length 2023;
Best Local Similarity 99.7%; Pred. No. 1.5e-157;
Matches 366; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 998 GAGCCTAAAGCTTTAAATTAATTTGATAGAGCTATTAAAGCTGAACCTGATGCTGCAGAG 1057
DB 998 GAGCCTAAAGCTTTAAATTAATTTGATAGAGCTATTAAAGCTGAACCTGATGCTGCAGAG 943
QY 1058 AAACTCCGTAAGGGGAAAAATTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGC 1117
DB 944 AAACTCCGTAAGGGGAAAAATTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGC 1003
QY 1118 CAATATACCTCAGAGTCCACGAGCAAGATATGGGAAGCGCAGTGTGAGGATGATTGGC 1177
DB 1004 CAATATACCTCAGAGTCCACGAGCAAGATATGGGAAGCGCAGTGTGAGGATGATTGGC 1063
QY 1178 TGAGAAGGAGAGAGTAAGTACGTGCTACGTGGAGCCATCGAGACCTCCAGAGGTGGC 1237
DB 1064 TGAGAAGGAGAGAGTAAGTACGTGCTACGTGGAGCCATCGAGACCTCCAGAGGTGGC 1123
QY 1238 CAGCTACCTGATGCTCCGTCGACAGCTGCTGAGCTGAGTTTGAAGCGTGCCTGACAGAG 1297
DB 1124 CAGCTACCTGATGCTCCGTCGACAGCTGCTGAGCTGAGTTTGAAGCGTGCCTGACAGAG 1183
QY 1298 GCAACAATTTCTAGGTCATATGAGAGGTTCCCTGCTTACCTCGACAGATTAGTTCAACT 1357
DB 1184 GCAACAATTTCTAGGTCATATGAGAGGTTCCCTGCTTACCTCGACAGATTAGTTCAACT 1243
QY 1358 ATTTCCC 1364
|||||

Db 1244 ATTTCCC 1250

RESULT 63
BX111251
LOCUS         Homo sapiens 484 bp mRNA linear EST 07-FBB-2003
DEFINITION    IMAGE:1876994, mRNA sequence.
ACCESSION     BX111251
VERSION       BX111251
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS       1 (bases 1 to 484)
              Ebert, U., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,
              Radelof, U., Schneider, D. and Korn, B.
TITLE         Human UnigeneSet - RZPD3
JOURNAL       Unpublished (2003)
COMMENT       Contact: Ina Rolfs
              RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
              Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
              RZPD; IMAGP998F034597.
              RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
              Human UnigeneSet - RZPD3 (RZPDLIB No.972)
              http://www.rzpd.de/CloneCards/cgi-
              bin/showlib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
              RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
              Heubnerweg 6, D-14059 Berlin, Germany
              Tel: +49 30 32639 101
              Fax: +49 30 32639 111
              www.rzpd.de
              This clone is available royalty-free from RZPD;
              contact RZPD (clone@rzpd.de) for further information. Seq primer:
              M13r, Primer sequence: TTTACACAGAAACAGCTATGAC.
FEATURES      Location/Qualifiers
              1..484
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="IMAGP998F034597 ; IMAGE:1876994"
                        /tissue_type="pooled human melanocyte, fetal heart, and
                        pregnant uterus"
                        /lab_host="DH10B"
                        /clone_lib="Soares_NHMPu_S1"
                        /notes="Organ: mixed (see below); Vector: pTT73D-Pac
                        (Pharmacia) with a modified polylinker; Site 1: Not I;
                        Site 2: Eco RI; Equal amounts of plasmid DNA from three
                        normalized libraries (melanocyte 2NbIM, pregnant uterus
                        NbIPU, and fetal heart NbHH19) were mixed, and ss circles
                        were made in vitro. Following HAP purification, this DNA
                        was used as tracer in a subtractive hybridization
                        reaction. The driver was PCR-amplified cDNAs from pools of
                        5,000 clones made from the same 3 libraries. The pools
                        consisted of I.M.A.G.E. clones 260232-265223,
                        340488-345479, and 484488-489479."
              ORIGIN
Query Match      13.5%; Score 313; DB 5; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.7e-156;
Matches 313; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 848 AACTGCTCCCTCCTGAGGATAATCTCTGTAGAGATTTCACAGTAAATTTGTAGAGAAAGTAAG 907
DB 1 AACTGCTCCCTCCTGAGGATAATCTCTGTAGAGATTTCACAGTAAATTTGTAGAGAAAGTAAG 60
QY 908 CATTTTTCCTGTGTGAGAACACGAGAGTACACCAAGAACAAATAGAAAACAGATGA 967
DB 61 CATTTTTCCTGTGTGAGAACACGAGAGTACACCAAGAACAAATAGAAAACAGATGA 120
QY 968 TCCAGACAAAAGACAAAGTTAGAAAAGAGAGCTTAAACTTTTAAATAATTTGATAA 1027
|||||
```

```

Db      121  TCAGAAACAAAAGCAGAAAGTTAAGAAAAAGAGCCCTAAACCTTTTAAATAAATTGATAA 180
Qy      1028 GACTATTAAAGCTGAACCTGATGCTGCAGAAAAAAGTCCGTAAGAGGGGAAAAATTGAGGA 1087
Db      181  GACTATTAAAGCTGAACCTGATGCTGCAGAAAAAAGTCCGTAAGAGGGGAAAAATTGAGGA 240
Qy      1088 AGCAGTGAATGCACTTTAAAGAACTAGTAGCAGCAATACCCCTCAGAGTCCAGCAGCAAGATA 1147
Db      241  AGCAGTGAATGCACTTTAAAGAACTAGTAGCAGCAATACCCCTCAGAGTCCAGCAGCAAGATA 300
Qy      1148 TGGGAAGCGCGAG 1160
Db      301  TGGGAAGCGCGAG 313

RESULT 64
CFI36513
LOCUS
DEFINITION UI-HF-BNO-anb-d-05-0-UI.r1 NIH MGC_50 Homo sapiens cDNA clone
IMAGE:3092001 5', mRNA sequence.
ACCESSION CFI36513
VERSION CFI36513.1 GI:33251957
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 503)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 9704477
PubMed 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Louis Staudt
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/humanfl.html
Seq primer: pYX-5,
Location/Qualifiers
FEATURES
source
1..503
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3092001"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/lab_host="DH10B (lTI)"
/clone_lib="NIH MGC 50"
/note="Vector: pYX3-Pac; Site 1: NotI; Site 2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(3.5-4.4kb). Directionally cloned. Cells provided by
Louis M. Staudt, Ph.D. Library preparation by Maria de
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

ORIGIN
Query Match 13.4%; Score 312; DB 7; Length 503;
Best Local Similarity 99.5%; Pred. No. 1.6e-155;
Matches 412; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      907 GCATTTTCTCTGGGAAGACAGCAGCAAGTACCAGAAACAAATAGAAAAACAGATG 966
Db      59 GCATTTTCTCTGGGAAGACAGCAGCAAGTACCAGAAACAAATAGAAAAACAGATG 118

```

```

Qy      967 ATCCGAAACAAAAGCAGAAAGTTAAGAAAAAGAGCCCTAAACCTTTTAAATAAATTGATA 1026
Db      119  ATCCGAAACAAAAGCAGAAAGTTAAGAAAAAGAGCCCTAAACCTTTTAAATAAATTGATA 178
Qy      1027 AGACTATTAAAGCTGAACCTGATGCTGCAGAAAAAAGTCCGTAAGAGGGGAAAAATTGAGG 1086
Db      179  AGACTATTAAAGCTGAACCTGATGCTGCAGAAAAAAGTCCGTAAGAGGGGAAAAATTGAGG 238
Qy      1087 AAGCAGTGAATGCACTTTAAAGAACTAGTAGCAGCAATACCCCTCAGAGTCCAGCAGCAAGAT 1146
Db      239  AAGCAGTGAATGCACTTTAAAGAACTAGTAGCAGCAATACCCCTCAGAGTCCAGCAGCAAGAT 298
Qy      1147 ATGGGAAGCGCGAGTGTGAGGATGATTTGGCTGAGAAGAGGAGCAAGTAATAGGTGCTAC 1206
Db      299  ATGGGAAGCGCGAGTGTGAGGATGATTTGGCTGAGAAGAGGAGCAAGTAATAGGTGCTAC 358
Qy      1207 GTGGAGCCATCGAGACTACCAAGAGTGGCAGCTTACCTGATGTCCTGCAGACCTGC 1266
Db      359  GTGGAGCCATCGAGACTACCAAGAGTGGCAGCTTACCTGATGTCCTGCAGACCTGC 418
Qy      1267 TGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAATTTCTAGTCTATGA 1320
Db      419  TGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAATTTCTAGTCTATGA 472

RESULT 65
BG170148
LOCUS
DEFINITION BG170148 682 bp mRNA linear EST 06-FEB-2001
IMAGE:4424894 5',
mRNA sequence.
ACCESSION BG170148
VERSION BG170148.1 GI:12676936
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 682)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10169 row: h column: 15
High quality sequence stop: 682.
Location/Qualifiers
FEATURES
source
1..682
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4424894"
/tissue_type="hypermphroma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 89"
/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.3 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 13.4%; Score 311; DB 4; Length 682;
Best Local Similarity 100.0%; Pred. No. 5.9e-155;
Matches 311; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```





## ORIGIN

Query Match	13.2%;	Score 306;	DB 7;	Length 560;
Best Local Similarity	99.7%;	Pred. No. 2.8e-152;		
Matches 356;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	CGGACCGTCGAATGGCCCGCAGCGTGAAGAATGCCAAGAGCAGCGCAACAGCAGCAGCAGCG	60	
Db	204	CGGACCGTCGAATGGCCCGCAGCGTGAAGATGCCAAGAGCAGCGCAACAGCAGCAGCAGCG	263	
QY	61	GCTCCGGCAGCGGTAGCAGAGTGCGGGAGCAGCAGCCCGGGGCCGGAGAGAGACAA	120	
Db	264	GCTCCGGCAGCGGTAGCAGAGTGCGGGAGCAGCAGCCCGGGGCCGGAGAGAGACAA	323	
QY	121	AGCATGGAGACACAGAATGGAGAGAAAGCGCGACTCTCGGGAATTCATTCTTCACGT	180	
Db	324	AGCATGGAGACACAGAATGGAGAGAAAGCGCGACTCTCAGGAATTCATTCTTCACGT	383	
QY	181	GGTTTATGGTGATTGCATTGCTCGGGCGTCGGACATCTGTAGTGTGCTTTGGTTTGATC	240	
Db	384	GGTTTATGGTGATTGCATTGCTCGGGCGTCGGACATCTGTAGTGTGCTTTGGTTTGATC	443	
QY	241	TTCTTGCACTATGAGGAAGTTCTAGGAAAACTAGGAAATCTATGATGCTGATGGTGATCGAG	300	
Db	444	TTGTTGACTATGAGGAAGTTCTAGGAAAACTAGGAAATCTATGATGCTGATGGTGATCGAG	503	
QY	301	ATTTTGATGGGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTTCAG	357	
Db	504	ATTTTGATGGGATGATGCCAAAGTTTATTAGGACTTAAAGAGAGATCTACTTTCAG	560	

RESULT	69
CN368828	
LOCUS	548 bp mRNA linear EST 16-MAY-2004
DEFINITION	17000424344539 GRN_EB Homo sapiens cDNA 5', mRNA sequence.
ACCESSION	CN368828
VERSION	CN368828.1 GI:47368762
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 548) Li, Y., Xu, C., Fang, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G. J., Brandenberger, R., Guegler, K., Rao, M. S., Mandalan, R., Lebkowski, J and Stanton, L.W Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation Nat. Biotechnol. 22 (6), 707-716 (2004)
AUTHORS	Contact: Brandenberger R Regenerative Medicine Geron Corporation 230 Constitution Drive, Menlo Park, CA 94025, USA Tel: 650 473 8658 Fax: 650 473 7760 Email: rbrandenberger@geron.com Insert Length: 548 Std Error: 0.00.
TITLE	
JOURNAL	
COMMENT	

**FEATURES**  
**SOURCE**

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, embryoid bodies
derived from H1, H7 and H9 cells"
/clone_lib="GRN EB"
/notes="oligo dt primed, full-length enriched cDNA library"
```

	Matches	408;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
Qy	153	GGACTCTCGGAACTTCATTC	CTTCACGTCGTTTATCGTG	AATTCGATTCGATTCGTCGGCTCTGG	212					
Db	125	GGACTCTCGGAACTTCATTC	CTTCACGTCGTTTATCGTG	AATTCGATTCGATTCGTCGGCTCTGG	184					
Qy	213	ACAATCTGAGCTGTCGTTTGG	TTTGATCTCTGTTGACTATCAG	GAAGTAGTTCCTAGGAAACTA	272					
Db	185	ACGTCGTGAGCTGTCGTTTGG	TTTGATCTCTGTTGACTATCAG	GAAGTAGTTCCTAGGAAACTA	244					
Qy	273	GGATCTATGATGCTGATGGTG	TGATGGAGATTTTGTGATGTG	GATGATGCCAAAGTTTATTATTA	332					
Db	245	GGATCTATGATGCTGATGGTG	TGATGGAGATTTTGTGATGTG	GATGATGCCAAAGTTTATTATTA	304					
Qy	333	GGACTTAAAGAGAGATCTACT	TTCAGACCCAGCAGTCCCGC	CAAGAGAGCTGAGCCACAC	392					
Db	305	GGACTTAAAGAGAGATCTACT	TTCAGAGCCAGCAGTCCCGC	CAAGAGAGCTGAGCCACAC	364					
Qy	393	ACTGAGCCCGAGGACAGGTTCT	GTGCGAGCCAGAACCCCGA	GAATATCGAAGATCAAGCA	452					
Db	365	ACTGAGCCCGAGGACAGGTTCT	GTGCGAGCCAGAACCCCGA	GAATATCGAAGATCAAGCA	424					
Qy	453	AAAGAACAAATTCAGTCCCTTC	TCATCATGTAATGGTACACG	CAACAATGTTTGAGGGAGAA	512					
Db	425	AAAGAACAAATTCAGTCCCTTC	TCATCATGTAATGGTACACG	CAACAATGTTTGAGGGAGAA	484					
Qy	513	GACTTTGCCAACAGAAAGATGG	ACCACACGAGAACCAACAAG	GAGATCA	562					
Db	485	GACTTTGCCAACAGAAAGATGG	ACCACACGAGAACCAACAAG	GAGATCA	534					

RESULT	68
CN482623	
LOCUS	
DEFINITION	CN483623 hw22f04.y1 Human primary ocular pericytes. Unamplified (hw) <i>Homo sapiens</i> cDNA clone hw22f04 5' mRNA sequence.
	EST 26-APR-2004

ACCESSION	CN482623	
VERSION	CN482623.1	GI:46564127
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1	(bases 1 to 560)
AUTHORS	Tsai,J.Y. and Wistow,G.	
TITLE	Expressed sequence tag analysis of cultured primary human ocular	
	pericytes	
JOURNAL	Unpublished (2004)	
COMMENT	Contact: Wistow G	
	Section on Molecular Structure and Function	
	National Eye Institute	
	6/331, NIH, Bethesda, MD 20892-2740, USA	
	Tel: 301 402 3452	
	Fax: 301 496 0078	
	Email: graeme@helix.nih.gov	
	Plate: 22	row: f column: 04
	Seq primer: M13p1	reverse primer (ABI).

```

FEATURES
source
Location/Qualifiers
1..560
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="hw22f04"
/cell_type="pericytes"
/dev_stage="Adult"
/lab_host="EMDH108"
/clone_lib="Human primary human ocular pericytes.
Unamplified (hw)"
/note="Organ: Eye; Vector: pSport1; RNA was extracted from
primary human pericytes in culture. A directionally cloned
cDNA library in the pSPORT1 vector (Invitrogen) was
constructed at Bioserve Biotechnology (Laurel MD)
essentially following the protocols of the SuperScript

```

from embryoid body outgrowths derived from hES cell lines  
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free  
conditions."

```

ORIGIN
Query Match      12.8%; Score 297; DB 7; Length 548;
Best Local Similarity 99.7%; Pred. No. 1.9e-147;
Matches 347; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGACCGTGCATGCGCCAGCGTAAGAAATGCCAAGAGCAGCGGCAACAGCAGCAGCG 60
DB 201 CGACCGTGCATGCGCCAGCGTAAGAAATGCCAAGAGCAGCGGCAACAGCAGCAGCG 260
QY 61 GTCGCGGAGCGGTAGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCG 120
DB 261 GTCGCGGAGCGGTAGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCGTGCAGCG 320
QY 121 AGCATGAGGACACAGAAATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
DB 321 AGCATGAGGACACAGAAATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 380
QY 181 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
DB 381 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 440
QY 241 TTGTTGACTATGAGAAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGCTGATG 300
DB 441 TTGTTGACTATGAGAAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGCTGATG 500
QY 301 ATTTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 348
DB 501 ATTTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 548

```

## RESULT 70

```

BE018154
LOCUS      509 bp      mRNA      linear      EST 06-JUN-2000
DEFINITION bb76a02.y1 NIH MGC 12 Homo sapiens cDNA clone IMAGE:3048266 5',
similar to SW:ASPH_HUMAN Q12797 ASPATYL/ASPARAGINYL
BETA-HYDROXYLASE ; mRNA sequence.
BE018154
VERSION    1
KEYWORDS   GI:8278168
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/resources.shtml
Seq primer: 5' 40RP from Gibco
High quality sequence stop: 447.
Location/Qualifiers
1..509
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3048266"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/notes="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.

```

## FEATURES

```

source
1..509
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3048266"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/notes="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.

```

Average insert size 1.4 kb. Library prepared by Life  
Technologies."

## ORIGIN

```

Query Match      12.7%; Score 295; DB 2; Length 509;
Best Local Similarity 99.7%; Pred. No. 2.3e-146;
Matches 345; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1979 GACACAGAGTGCAGAGAGGACAGATCAATATATTCATCATGACACCCCGGACTCACGT 2038
DB 1 GACACAGAGTGCAGAGAGGACAGATCAATATATTCATCATGACACCCCGGACTCACGT 60
QY 2039 GTGCGCCGACACAGAGGCGCCACAACTGCGAGGCTCCGATGCGACCTGGGCTTGTGATTC 2098
DB 61 GTGCGCCGACACAGAGGCGCCACAACTGCGAGGCTCCGATGCGACCTGGGCTTGTGATTC 120
QY 2099 CAGAGAGGCTGCAGAGATTCGATGTGCCAAGAGCAGGACCTGGAGGAGGAGGAGGT 2158
DB 121 CAGAGAGGCTGCAGAGATTCGATGTGCCAAGAGCAGGACCTGGAGGAGGAGGAGGT 180
QY 2159 GCTCATCTTTGATGACTCTCTTTGAGCAGAGGATGCGAGGATGCGCTCATCTTTCCGGCT 2218
DB 181 GCTCATCTTTGATGACTCTCTTTGAGCAGAGGATGCGAGGATGCGCTCATCTTTCCGGCT 240
QY 2219 GATATTCATGCTGATGCTGCGCATCCGAACTGACACCAAGAGGAGGAGGAGGAGGAG 2278
DB 241 GATATTCATGCTGATGCTGCGCATCCGAACTGACACCAAGAGGAGGAGGAGGAGGAG 300
QY 2279 AGCAATTTAGCATGAATTCATGCAAGCTTGGGAAACTCTGGAGAGA 2324
DB 301 AGCAATTTAGCATGAATTCATGCAAGCTTGGGAAACTCTGGAGAGA 346

```

## RESULT 71

```

BE0785890
LOCUS      536 bp      mRNA      linear      EST 05-MAR-2002
DEFINITION K-EST0064453 S9SNU601 Homo sapiens cDNA clone S9SNU601-34-C04 5',
mRNA sequence.
BE0785890
VERSION    1
KEYWORDS   GI:19134122
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 536)
Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and
Kim, Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongseung@mail.kribb.re.kr
Plate: 34 row: C column: 04
High quality sequence stop: 536.
Location/Qualifiers
1..536
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S9SNU601-34-C04"
/sex="M"
/tissue_type="Ascites"
/cell_type="Epithelial"
/cell_line="SNU-601"
/lab_host="Top10F"
/clone_lib="S9SNU601"
/notes="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;

```

## FEATURES

```

source
1..536
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S9SNU601-34-C04"
/sex="M"
/tissue_type="Ascites"
/cell_type="Epithelial"
/cell_line="SNU-601"
/lab_host="Top10F"
/clone_lib="S9SNU601"
/notes="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;

```

Site 2: XhoI: The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including SfiI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized with Superscript II using SfiI oligo-dT primer. After first strand synthesis, RNA was degraded by NaOH treatment and cDNA was amplified by PCR reaction. The PCR products were digested with SfiI and cloned into DraIII- digested pME18S-FL3 vector. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 12.7%; Score 294; DB 4; Length 536;  
Best Local Similarity 100.0%; Pred. No. 8e-146;  
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 508 GAGAAGACTTCCAAACAGAAGATGACCCACAGGAGAACACACACAGAGGATGATGAGT 567  
Db 1 GAGAAGACTTCCAAACAGAAGATGACCCACAGGAGAACACACACAGAGGATGATGAGT 60

Qy 568 TTCTTATGGGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 627  
Db 61 TTCTTATGGGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 120

Qy 628 AAGAACCAGCAGATAGTTACCACTGGGAAGAGACAGTTTCAACAGCTGTTAATCAGGATA 687  
Db 121 AAGAACCAGCAGATAGTTACCACTGGGAAGAGACAGTTTCAACAGCTGTTAATCAGGATA 180

Qy 688 TCGAAGACAGATGCTGACGAGGAGGAATCCAGATCCAGTGAACCCAGTGTAGAGATG 747  
Db 181 TCGAAGACAGATGCTGACGAGGAGGAATCCAGATCCAGTGAACCCAGTGTAGAGATG 240

Qy 748 AAGAGTTGCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 801  
Db 241 AAGAGTTGCACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 294

RESULT 72

AV751514 622 bp mRNA linear EST 19-OCT-2000  
LOCUS AV751514 NPD Homo sapiens cDNA clone NPDXB01 5', mRNA sequence.

DEFINITION AV751514  
ACCESSION AV751514 GI:10909362  
VERSION EST.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 622)  
AUTHORS Song,H., Peng,Y., Gu,Y., Yang,Y., Gao,G., Xiao,H., Xu,X., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Ye,M., Zhang,Q., Han,Z., Chen,Z., Hu,R. and Chen,J.  
JOURNAL Homo sapiens NPD library cDNA clones  
COMMENT Unpublished (2000)  
Contact: Qinghua Zhang  
Shanghai Institute of Endocrinology, Rui-Jin Hospital  
197 Rui-Jin II Road, Shanghai 200025, P. R. China  
Tel: 86-21-64370045 (ex.663332)  
Fax: 86-21-64743206  
Email: mshn@ms.sh.cn  
This clone is available at Shanghai Hematology Institute in Shanghai.  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi- Tech Park, Pudong.  
Location/Qualifiers  
1. .622  
/organism="Homo sapiens"  
/mol\_type="mRNA"

FEATURES  
source

/db\_xref="taxon:9606"  
/clone="NPDXB01"  
/tissue\_type="pituitary"  
/dev\_stage="Adult"  
/lab\_host="SOLB"  
/clone\_lib="NPD"  
/note="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2: XhoI"

ORIGIN

Query Match 12.7%; Score 294; DB 1; Length 622;  
Best Local Similarity 100.0%; Pred. No. 8.2e-146;  
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 653 GGAAGAGACAGATTTTCACAAAGACTGTAATCAGGATATGGAAGAGATGATCTTGAGCAGGA 712  
Db 90 GGAAGAGACAGATTTTCACAAAGACTGTAATCAGGATATGGAAGAGATGATCTCTGAGCAGGA 149

Qy 713 AAATCCAGATTCAGTGAACCCAGTAGTAGAAGATGAAGATTCACCATGATACAGATGA 772  
Db 150 AAATCCAGATTCAGTGAACCCAGTAGTAGAAGATGAAGATTCACCATGATACAGATGA 209

Qy 773 TGTAAACATACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAAAATGAAGGGAT 832  
Db 210 TGTAAACATACCAAGTCTATGAGGAACAAGCAGTATATGAACCTCTAGAAAAATGAAGGGAT 269

Qy 833 AGAAATCAGACAGAGTAGTAACCTGCTCCCTCAGGATATATCTGTAGAAGATTCACAGTAAT 892  
Db 270 AGAAATCAGACAGAGTAGTAACCTGCTCCCTCAGGATATATCTGTAGAAGATTCACAGTAAT 329

Qy 893 TGTAGAAGAGTAGAAGATTTTCTCTGGAAGAACAGCAGGAGGAGTACCACCCAGA 946  
Db 330 TGTAGAAGAGTAGAAGATTTTCTCTGGAAGAACAGCAGGAGGAGTACCACCCAGA 383

RESULT 73

CD512083 862 bp mRNA linear EST 06-JUN-2003  
LOCUS AGENCOURT\_14354806 NIH MGC\_186 Homo sapiens cDNA clone  
DEFINITION IMAGE:30406369 5', mRNA sequence.

ACCESSION CD512083  
VERSION CD512083.1 GI:31443801  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 862)  
AUTHORS NIH-MGC http://mgs.nci.nih.gov/  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cagpos-remail.nih.gov  
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: NDCMI95 row: a column: 02  
High quality sequence stop: 564.  
Location/Qualifiers  
1. .862  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30406369"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH\_MGC\_186"

FEATURES  
source

/note="Organ: Pooled-Skin; Vector: pDNR-LIB; Site 1: SfiI (ggccatagcc); Site 2: SfiI (ggcgctggcc); Library is oligo-dT primed and directionally cloned. cDNA was prepared from a pooled samples of tissues from Skin, meninges, duramater, pia mater and choroid plexus. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGCC-3' and 3' adaptor sequence: 5'-ATTCTAGCGCGAGCGCCACATG-T(30)BN-3' (where B = A, C, G or T). Average insert size 1.47 kb (range 0.50-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library"

## ORIGIN

Query Match 12.7%; Score 294; DB 6; Length 862;  
Best Local Similarity 100.0%; Pred. No. 8.7e-146; Indels 0; Gaps 0;  
Matches 294; Conservative 0; Mismatches 0;

QY 653 GGAAGAGACAGTTTCACAGACTGTATCAGGATATGGAAGAGATGATGTCTGAGCAGGA 712  
DB 3 GGAAGAGACAGTTTCACAGACTGTATCAGGATATGGAAGAGATGATGTCTGAGCAGGA 62  
QY 713 AAATCCAGATTCAGTGAACACAGTATGAAAGATGCAACATGATACAGATGA 772  
DB 63 AAATCCAGATTCAGTGAACACAGTATGAAAGATGCAACATGATACAGATGA 122  
QY 773 TGTACATACCAAGTCTATGAGGAACAAGCAGTATATCACTCTAGAAATCAAGGAT 832  
DB 123 TGTACATACCAAGTCTATGAGGAACAAGCAGTATATCACTCTAGAAATCAAGGAT 182  
QY 833 AGAATACACAGTACTCTCTCCCTCAGGATATCTCTAGAGATTCACAGTAAAT 892  
DB 183 AGAATACACAGTACTCTCTCCCTCAGGATATCTCTAGAGATTCACAGTAAAT 242  
QY 893 TGTAGAAGATTAAGCATTTTCTCTGTGAAGAACAGCAGGAAGTACCACAGA 946  
DB 243 TGTAGAAGATTAAGCATTTTCTCTGTGAAGAACAGCAGGAAGTACCACAGA 296

## RESULT 74

BM462707  
LOCUS BM462707  
DEFINITION AGENCOURT\_6427239 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:5540329  
5', mRNA sequence.

ACCESSION BM462707  
VERSION BM462707.1 GI:18511747

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1010)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbe-remail.nih.gov](mailto:cgapbe-remail.nih.gov)

Tissue Procurement: ATCC/DCTD/DTT

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Cloning Distribution: Agencourt Bioscience Corporation

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LAM12236 row: e column: 02

High quality sequence stop: 579.

Location/Qualifiers

1..1010

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

## FEATURES

source

/clone="IMAGE:5540329"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_72"  
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2 kb. Library constructed by Life  
Technologies."

## ORIGIN

Query Match 12.4%; Score 288; DB 4; Length 1010;  
Best Local Similarity 99.5%; Pred. No. 1.5e-142;  
Matches 388; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGACCGTGCATATGCCCGCAGCGTATGAAATGCCAAGACGCGCGCACAGCAGCAGCG 60  
DB 206 CGGACCGTGCATATGCCCGCAGCGTATGAAATGCCAAGACGCGCGCACAGCAGCAGCG 265  
QY 61 GCTCCGCGAGCGGTAGCAGAGTGGCGGCGAGCAGCAGCCCGCGGCGGAGAGACAA 120  
DB 266 GCTCCGCGAGCGGTAGCAGAGTGGCGGCGAGCAGCAGCCCGCGGCGGAGAGACAA 325  
QY 121 AGCATGGAGGACACAAAGATGGGAGGAAAGCGGACTCTCGGAACTTCATCTTCACGT 180  
DB 326 AGCATGGAGGACACAAAGATGGGAGGAAAGCGGACTCTCGGAACTTCATCTTCACGT 385  
QY 181 GGTATTATGATGATTCATTTGCTGGGCGCTCTGGACATCTGATGCTGCTGTTGTTGATC 240  
DB 386 GGTATTATGATGATTCATTTGCTGGGCGCTCTGGACATCTGATGCTGCTGTTGTTGATC 445  
QY 241 TTGTTGATATGAGGAAGTTCTAGGAAACTATGAAATCTATGATGCTGATGATGAG 300  
DB 446 TTGTTGATATGAGGAAGTTCTAGGAAACTATGAAATCTATGATGCTGATGATGAG 505  
QY 301 ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360  
DB 506 ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 565  
QY 361 CAGCAGTCCCGCCAGAGAGAGGCTGAGCCAC 390  
DB 566 CAGCAGTCCCGCCAGAGAGAGGCTGAGCCAC 595

## RESULT 75

CN368820  
LOCUS CN368820  
DEFINITION 17000533684042 GRN\_EB Homo sapiens cDNA 5', mRNA linear EST 16-MAY-2004  
ACCESSION CN368820  
VERSION CN368820.1 GI:47368754

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 767)

Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,

Li, Y., Xu, C., Pang, R., Guelet, K., Rao, M.S., Mandalam, R.,

Lebkowski, J. and Stanton, L.W.

Transcriptome characterization elucidates signaling networks that

control human ES cell growth and differentiation

Nat. Biotechnol. 22 (6), 707-716 (2004)

Contact: Brandenberger R

Regenerative Medicine

Genon Corporation

230 Constitution Drive, Menlo Park, CA 94025, USA

Tel: 650 473 8658

Fax: 650 473 7760

Email: [rbrandenberger@genon.com](mailto:rbrandenberger@genon.com)

Insert Length: 767 Std Error: 0.00.

Location/Qualifiers

1..767

/organism="Homo sapiens"

/mol\_type="mRNA"

/db xref="taxon:9606"  
/tissue type="embryonic stem cells, embryoid bodies  
derived from H1, H7 and H9 cells"  
/clone\_lib="GRN\_EB"  
/notes="oligo dt primed, full-length enriched cDNA library  
from embryoid body outgrowths derived from hES cell lines  
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free  
conditions."

ORIGIN

Query Match 12.2%; Score 283; DB 7; Length 767;  
Best Local Similarity 99.7%; Pred. No. 7.1e-140;  
Matches 333; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CGGACCGTGCATCGCCAGCGTAAGATGCCAAGAGCAGCGCAACAGCAGCAGCAGCG 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
180 CGGACCGTGCATCGCCAGCGTAAGATGCCAAGAGCAGCGCAACAGCAGCAGCAGCG 239  
Qy 61 GCTCCGGCAGCGGTAGCACGAGTGGGGCAGCAGCAGCCCGGGGCCCGGAGAGACAA 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
240 GCTCCGGCAGCGGTAGCACGAGTGGGGCAGCAGCAGCCCGGGGCCCGGAGAGACAA 299  
Qy 121 AGCATGGAGGACACAAGAAATGGAGGAAAGCCGACTCTCGGAACTTCATTCTTACGT 180  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
300 AGCATGGAGGACACAAGAAATGGAGGAAAGCCGACTCTCAGGAACTTCATTCTTACGT 359  
Qy 181 GGTATTATGTTGATTCGATTGCTGGCGCTCTGGACATCTGTAGCTGTGTTGGTTGATC 240  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
360 GGTATTATGTTGATTCGATTGCTGGCGCTCTGGACATCTGTAGCTGTGTTGGTTGATC 419  
Qy 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGGTATGGAG 300  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
420 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGGTATGGAG 479  
Qy 301 ATTTTGATGTGGATGATCCCAAGTTTATTAGG 334  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
480 ATTTTGATGTGGATGATCCCAAGTTTATTAGG 513

Search completed: March 25, 2005, 22:28:27  
Job time : 6796 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 14:29:03 ; Search time 1157 Seconds  
(without alignments)  
11890.647 Million cell updates/sec

Title: US-09-436-184-3  
Perfect score: 2324  
Sequence: 1 cggaccgtgcaatggccag.....cttgggaactctggagaga 2324

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

N\_Geneseq\_16Dec04.\*  
1: Geneseqn1980s.\*  
2: Geneseqn1980s.\*  
3: Geneseqn2000as.\*  
4: Geneseqn2000as.\*  
5: Geneseqn2000as.\*  
6: Geneseqn2000as.\*  
7: Geneseqn2000as.\*  
8: Geneseqn2000as.\*  
9: Geneseqn2000as.\*  
10: Geneseqn2000as.\*  
11: Geneseqn2000as.\*  
12: Geneseqn2000as.\*  
13: Geneseqn2000as.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2324	100.0	2324	4	Aaf89811 cDNA enco
2	2324	100.0	2324	8	Abs57868 Human cDN
3	2324	100.0	2324	9	Ada00640 cDNA enco
4	2322	95.6	2452	13	Adp23914 PRO polyp
5	2322	95.6	2452	13	Adp23914 PRO polyp
6	2171	93.4	5358	10	Adj56257 Bovine CD
7	895	38.5	2680	6	Abx04178 Human mRN
8	895	38.5	2680	13	Adp24208
9	840	27.5	2442	2	Aaz23609 Human lab
10	840	27.5	2442	8	Abk39743 cDNA enco
11	840	27.5	2442	6	Acad2072 Human lunc
12	840	27.5	2442	8	Acad2072 Human lunc
13	640	27.5	2442	10	Adh47310 Human lun
14	640	27.5	2442	13	Adj21229 Human lun
15	615	26.5	3110	11	Acn88788 Breast ca
16	502	21.6	2648	10	Acn88788 Breast ca
17	459	19.8	502	9	Ach47067 Human inf
18	359	15.4	660	6	Abq59507 Human col
19	249	10.7	495	11	Adt94758
20	104	4.5	161	12	Ach93769 Human gen

94 20 0.9 28564 10 ADD45153 Add45153 Human gen  
 95 20 0.9 48037 4 AAK84729 Aak84729 Human imm  
 96 20 0.9 48037 4 AAK85983 Aak85983 Human imm  
 97 20 0.9 48045 4 AAK84730 Aak84730 Human imm  
 98 20 0.9 48045 4 AAK85984 Aak85984 Human imm  
 c 99 20 0.9 52677 11 ACN43934 Acn43934 Human gen  
 c 100 20 0.9 159400 6 ABQ88126 Abq88126 Human ost

## ALIGNMENTS

## RESULT 1

AA89811  
 ID AA89811 standard; cDNA; 2324 BP.  
 XX  
 AC AAF89811;  
 XX  
 DT 23-JUL-2001 (first entry)  
 XX  
 DE cDNA encoding human aspartyl (asparaginyl) beta-hydroxylase (HAAH).  
 XX  
 KW Epidermal growth factor-like domain; EGF-like domain; cancer;  
 KW human aspartyl beta-hydroxylase; HAAH; malignant neoplasm; tumour; ss.  
 XX  
 OS Homo sapiens.

## Location/Qualifiers

XX Key 12..2289  
 XX CDS /\*tag= a  
 FT /product= "human aspartyl (asparaginyl) beta-hydroxylase"  
 FT  
 XX

PN WO200135102-A2.

XX  
 PD 17-MAY-2001.

XX 08-NOV-2000; 2000WO-US030738.

XX 08-NOV-1999; 99US-00436184.

XX (RHOD-) RHODE ISLAND HOSPITAL LIFESPAN PARTNER.

XX Wands JR, De La Monte SM, Ince N, Carlson RI;

XX WPI; 2001-329171/34.

XX P-PsDB; AAB83919.

XX Diagnosing malignant neoplasm in a mammal, involves contacting mammalian  
 PT sample with antibody that binds to human aspartyl beta-hydroxylase  
 PT polypeptide to form antigen-antibody complex and detecting the complex.

XX Disclosure; Page 6-7; 76pp; English.

XX The present sequence encodes a human aspartyl (asparaginyl) beta-  
 CC hydroxylase (HAAH) enzyme. Epidermal growth factor (EGF)-like domains of  
 CC polypeptides are hydroxylated by HAAH enzymes. HAAH is used in the method  
 CC of the invention. The specification describes a method for diagnosing a  
 CC malignant neoplasm in a mammal. The method comprises contacting a body  
 CC fluid with an antibody which binds to HAAH polypeptide under complex  
 CC forming conditions, and detecting the antigen-antibody complex. The  
 CC method is useful for diagnosing and prognosing a malignant neoplasm in a  
 CC bodily fluid e.g. central nervous system (CNS)-derived body fluid, blood,  
 CC serum, urine, saliva, sputum, lung effusion, and ascites fluid of mammal,  
 CC where the neoplasm is derived from endodermal tissue and is selected from  
 CC colon cancer, breast cancer, pancreatic cancer, liver cancer, cancer of  
 CC bile ducts, and cancer of the CNS. HAAH antibodies, linked to a cytotoxic  
 CC agent, are useful for killing tumour cells

SQ Sequence 2324 BP; 753 A; 463 C; 625 G; 483 T; 0 U; 0 Other;

Query Match 100.0%; Score 2324; DB 4; Length 2324;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCGTGCATATGGCCAGCGTAAGAATCCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
 DB 1 CGGACCGTGCATATGGCCAGCGTAAGAATCCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
 QY 61 GCTCCGGCAGCGGTAGCAGCAGATGCGGGCAGCAGCAGCCCGGGCCCGGAGAGACAA 120  
 DB 61 GCTCCGGCAGCGGTAGCAGCAGATGCGGGCAGCAGCAGCCCGGGCCCGGAGAGACAA 120  
 QY 121 AGCATGGAGGACACAGAATGGGAGGAAAGCGGAGCTCTCGGGAACCTTCATTCTTACGCT 180  
 DB 121 AGCATGGAGGACACAGAATGGGAGGAAAGCGGAGCTCTCGGGAACCTTCATTCTTACGCT 180  
 QY 181 GGTATTATGTTGATTCCTGGCGTCTGGCACTCTGGAATCTATGATGCTGATGCTGATGATG 240  
 DB 181 GGTATTATGTTGATTCCTGGCGTCTGGCACTCTGGAATCTATGATGCTGATGCTGATGATG 240  
 QY 241 TTGTTGATATGAGGAAGTTCTAGGAAATCTAGGAATCTATGATGCTGATGCTGATGATGAG 300  
 DB 241 TTGTTGATATGAGGAAGTTCTAGGAAATCTAGGAATCTATGATGCTGATGCTGATGATGAG 300  
 QY 301 ATTTTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
 DB 301 ATTTTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
 QY 361 CAGCAGTCCCAGGAGAGAGGCTGAGCCACACTGAGCCCGGAGGAGCAGGTTCTCTGTTGG 420  
 DB 361 CAGCAGTCCCAGGAGAGAGGCTGAGCCACACTGAGCCCGGAGGAGCAGGTTCTCTGTTGG 420  
 QY 421 AGGAGAGACCCAGATATCGAAGATGAGCAAGAGCAAAATTCAGTCCCTCTTCCATG 480  
 DB 421 AGGAGAGACCCAGATATCGAAGATGAGCAAGAGCAAAATTCAGTCCCTCTTCCATG 480  
 QY 481 AAATGGTACACGAGAACTGTTGAGGAGAGAACTTGCACAAAGAGAGATGGACCCACAG 540  
 DB 481 AAATGGTACACGAGAACTGTTGAGGAGAGAACTTGCACAAAGAGAGATGGACCCACAG 540  
 QY 541 GAGAACCAACAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
 DB 541 GAGAACCAACAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
 QY 601 AGACCTGGACCTGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTGGAAGAGA 660  
 DB 601 AGACCTGGACCTGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTGGAAGAGA 660  
 QY 661 CAGTTTCAAGACTGTAAATCAGGATATGGAAGAGATGATGATGATGATGATGATGATGATGATG 720  
 DB 661 CAGTTTCAAGACTGTAAATCAGGATATGGAAGAGATGATGATGATGATGATGATGATGATGATG 720  
 QY 721 ATTCCAGTGAACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780  
 DB 721 ATTCCAGTGAACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 780  
 QY 781 ACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAAATGAAGGGATAGAAAATCA 840  
 DB 781 ACCAAGTCTATGAGGAAACAGCAGTATATGAACCTCTAGAAAATGAAGGGATAGAAAATCA 840  
 QY 841 CAGAAGTAACTGTCTCCCTGAGGATATCTCTGAGGATATCTCTGAGCAGGAAAATCCAG 900  
 DB 841 CAGAAGTAACTGTCTCCCTGAGGATATCTCTGAGGATATCTCTGAGCAGGAAAATCCAG 900  
 QY 901 AAGTAAAGCAATTTCTCTGAGGATATCTCTGAGGATATCTCTGAGCAGGAAAATCCAG 960  
 DB 901 AAGTAAAGCAATTTCTCTGAGGATATCTCTGAGGATATCTCTGAGCAGGAAAATCCAG 960  
 QY 961 CAGATGATCCAGAAACAAAAGTTAAGAAAAGAGCCTTAAATTAAT 1020  
 DB 961 CAGATGATCCAGAAACAAAAGTTAAGAAAAGAGCCTTAAATTAAT 1020  
 QY 1021 TTGATAAGACTATTAAAGCTGAACTTTGATGCTGCAGAAAAAATCCGTTAAAGGGGAAAAA 1080  
 DB 1021 TTGATAAGACTATTAAAGCTGAACTTTGATGCTGCAGAAAAAATCCGTTAAAGGGGAAAAA 1080



1081 TTGAGGAGCAGTGAATCAITTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140  
1081 TTGAGGAGCAGTGAATCAITTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140  
1141 CAAGATATGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGGAGGAGTATGAGG 1200  
1141 CAAGATATGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGGAGGAGTATGAGG 1200  
1201 TGCTAGCTGGACCATGAGACCTACCAAGAGGTGGCCAGCCTACCTGATGTCCTGGCAG 1260  
1201 TGCTAGCTGGAGCCATGAGACCTACCAAGAGGTGGCCAGCCTACCTGATGTCCTGGCAG 1260  
1261 ACTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAATTTCTAGGTCATATGA 1320  
1261 ACTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAATTTCTAGGTCATATGA 1320  
1321 GAGGTTCCCTGCTTACCTGCGAGAGATTAGTTCAACTATTTCCCAATGATACTTCTTAA 1380  
1321 GAGGTTCCCTGCTTACCTGCGAGAGATTAGTTCAACTATTTCCCAATGATACTTCTTAA 1380  
1381 AAATGACCTTGGCGTGGATACCTCTGATAGGAGATTAATGACAAATGCAAGAAAGTTT 1440  
1381 AAATGACCTTGGCGTGGATACCTCTGATAGGAGATTAATGACAAATGCAAGAAAGTTT 1440  
1441 ATGAAGAGTGTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTATGGCTTCA 1500  
1441 ATGAAGAGTGTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTATGGCTTCA 1500  
1501 TCCTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCCATATTTAAAGCAAGGAATAGAAT 1560  
1501 TCCTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCCATATTTAAAGCAAGGAATAGAAT 1560  
1561 CCGGAGATCTGGCACTGATGATGGAGATTTTATTTCCACCTGGGGGATGCCATCGAGA 1620  
1561 CCGGAGATCTGGCACTGATGATGGAGATTTTATTTCCACCTGGGGGATGCCATCGAGA 1620  
1621 GGGTTGGGAAACAAAGAGGCATATAAGTGTATGAGCTTTGGGCAACAGAGAGGACATTTG 1680  
1621 GGGTTGGGAAACAAAGAGGCATATAAGTGTATGAGCTTTGGGCAACAGAGAGGACATTTG 1680  
1681 CATCTGTCTGCAACCGCTCACTCAATGTGAATGGAATGGAATGGAATGGAATGGAATGGA 1740  
1681 CATCTGTCTGCAACCGCTCACTCAATGTGAATGGAATGGAATGGAATGGAATGGAATGGA 1740  
1741 CCCCAAAAGAAACGGGCTACACAGAGTGTAGTAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
1741 CCCCAAAAGAAACGGGCTACACAGAGTGTAGTAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
1801 TCCGAGATGAAGGCTTGCAGTGTGATGAATGAATGAATGAATGAATGAATGAATGAATGA 1860  
1801 TCCGAGATGAAGGCTTGCAGTGTGATGAATGAATGAATGAATGAATGAATGAATGAATGA 1860  
1861 AAACCTGAGGAAAGGGGACTGAGCCAGTTTCACTGTGTGGCAGCAAGGAAGAA 1920  
1861 AAACCTGAGGAAAGGGGACTGAGCCAGTTTCACTGTGTGGCAGCAAGGAAGAA 1920  
1921 ATGAAATGCTGCAAGAGGAGCTCCTTAAACCTGTACCTTACTAGAAAAGTTTCCCGAGA 1980  
1921 ATGAAATGCTGCAAGAGGAGCTCCTTAAACCTGTACCTTACTAGAAAAGTTTCCCGAGA 1980  
1981 CAACAGGATGAGAGGAGGACAGATCAATATTTCCATGTCACCCCGGAGTCACTGCT 2040  
1981 CAACAGGATGAGAGGAGGACAGATCAATATTTCCATGTCACCCCGGAGTCACTGCT 2040  
2041 GGCCGCAACAGGGCCCAAACTGAGGCTCCGAATGACCTTGGGCTTGGTGTATCCCA 2100  
2041 GGCCGCAACAGGGCCCAAACTGAGGCTCCGAATGACCTTGGGCTTGGTGTATCCCA 2100  
2101 AGGAAGGCTGCAAGATTCGATGTGCAACGAGACAGGACCTGGGAGGAGCAAGGTGC 2160  
2101 AGGAAGGCTGCAAGATTCGATGTGCAACGAGACAGGACCTGGGAGGAGCAAGGTGC 2160  
2161 TCATCTTTGATGACTCTTTTGAGCAGGAGTATGGCAGGATGCTCATCTTCTCCGCTGA 2220

Db 2161 TCATCTTTGATGACTCTTTTGAGCAGGATATGGCAGATGCTCATCTTCTCCGCTGA 2220  
Qy 2221 TATTTCATCGTGGATGTGTGGCATCCGGAATCTGACACCAACAGAGAGCGGCTTCCAG 2280  
Db 2221 TATTTCATCGTGGATGTGTGGCATCCGGAATCTGACACCAACAGAGAGCGGCTTCCAG 2280  
Qy 2281 CAATTTAGCATGAATTCATCAAGCTTTGGGAACTCTGGAGAGA 2324  
Db 2281 CAATTTAGCATGAATTCATCAAGCTTTGGGAACTCTGGAGAGA 2324

RESULT 2  
ABS57868  
ID ABS57868 standard; CDNA; 2324 BP.  
XX  
AC ABS57868;  
XX  
DT 06-FEB-2003 (first entry)  
XX  
DE Human cDNA encoding aspartyl (asparaginyl) beta-hydroxylase, HAAH.  
XX  
KW Human; ss; gene; aspartyl (asparaginyl) beta-hydroxylase; HAAH;  
KW cytoostatic; immunostimulant; antibody; neoplasm; tumour; PB50; 5C7;  
KW 19B; brain tumour; glioma; glioblastoma; astrocytoma; haemangioma;  
KW pancreatic carcinoma; colon cancer; breast cancer; pancreatic cancer;  
KW liver cancer; cancer of the bile ducts; primary malignant CNS neoplasm;  
KW metastatic CNS neoplasm.  
XX  
OS Homo sapiens.  
XX  
PH Key Location/Qualifiers  
FT 12..2288  
FT /\*tag= a  
FT /product= "HAAH"  
XX  
XX US2002110559-A1.  
XX  
XX 15-AUG-2002.  
XX  
XX 17-MAY-2001; 2001US-00859604.  
XX  
XX 08-NOV-1999; 99US-00436184.  
XX  
XX (WAND/) WANDS J R.  
XX (DMON/) DE LA MONTE S M.  
XX (DEUT/) DEUTCH A H.  
XX (GHAN/) GHANBARI H A.  
XX  
XX Wands JR, De La Monte SM, Deutch AH, Ghanbari HA;  
XX  
XX WPI; 2003-066676/06.  
XX P-PSDB; ABG72365.  
XX  
XX Diagnosing neoplasm and inhibiting tumor growth, by contacting tissue of  
XX mammal with detectably-labeled antibody which binds to human aspartyl  
XX (asparaginyl) beta-hydroxylase.  
XX  
XX Example 1; Page 4-5; 34pp; English.  
XX  
XX The invention relates to diagnosing a neoplasm and inhibiting tumour  
XX growth in a mammal, using an antibody that binds to human aspartyl  
XX (asparaginyl) beta-hydroxylase (HAAH). Diagnosing a neoplasm comprises  
XX contacting a tissue with a detectably-labelled antibody where an increase  
XX in level of antibody binding at tissue site compared to the level of  
XX binding to normal non-neoplastic tissue indicates the presence of a  
XX neoplasm at the tissue site. Inhibiting tumour growth in mammal involves  
XX administering the antibody conjugated to a cytotoxic agent to a mammal.  
XX Also included are a method of conferring an immune response to a tumour  
XX cell in a mammal, by administering the antibody, a method of inducing an  
XX HAAH-specific immune response in a mammal, by administering to the mammal  
XX an HAAH polypeptide (or a polynucleotide composition encoding the  
XX polypeptide, or its degenerate variant), a fragment of HAAH comprising an

CC extracellular domain and lacking a cytoplasmic domain of HAAH, an  
CC antibody or its fragment which binds to HAAH (where the antibody is FB50,  
CC 86A, 5C7 or 19B), a hybridoma cell line chosen from hybridoma FB50,  
CC HA386A, HA15CA and HA219B, and a fragment of HAAH which lacks enzymatic  
CC activity or alpha-ketoglutarate binding domain and epidermal growth  
CC factor (EGF)-like domain. The methods are useful for diagnosing neoplasm  
CC in a mammal, inhibiting tumour growth in a mammal, conferring an immune  
CC response to a brain tumour cell (e.g. glioma, glioblastoma, astrocytoma  
CC or haemangioma) in a mammal, for conferring immune response to a  
CC pancreatic carcinoma cell and for inducing a HAAH-specific immune  
CC response in a mammal. The method is useful for diagnosing malignant  
CC neoplasms derived from endodermal tissue, e.g. colon cancer, breast  
CC cancer, pancreatic cancer, liver cancer and cancer of the bile ducts. The  
CC method is also useful for diagnosing neoplasms of central nervous system  
CC (CNS) e.g. primary malignant CNS neoplasms, and for diagnosing brain  
CC cell origin and metastatic CNS neoplasms, and for diagnosing brain  
CC tumours e.g. glioma, glioblastoma, astrocytoma or haemangioma. The  
CC present sequence encodes human HAAH  
XX  
SQ Sequence 2324 BP; 753 A; 463 C; 625 G; 483 T; 0 U; 0 Other;  
Query Match 100.0%; Score 2324; DB 8; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGACCGTCAATGCCCGCAGCGTAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
DB 1 CGGACCGTGCATATGCCCGCAGCGTAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
QY 61 GCTCCGGCAGCGGTAGCAGAGTGGCGGCAGCAGCAGCGCGCGCGCGCGGAGAGACAA 120  
DB 61 GCTCCGGCAGCGGTAGCAGAGTGGCGGCAGCAGCAGCGCGCGCGCGCGGAGAGACAA 120  
QY 121 AGCATGGAGGACAAGAATGGGAGGAAGGGGACTCTCGGGAACCTTCATCTTCACGT 180  
DB 121 AGCATGGAGGACAAGAATGGGAGGAAGGGGACTCTCGGGAACCTTCATCTTCACGT 180  
QY 181 GGTATTATGGTATTCATCTGCTGGCGCTCTGGACATCTGTAGCTGTCTGTTTGTATC 240  
DB 181 GGTATTATGGTATTCATCTGCTGGCGCTCTGGACATCTGTAGCTGTCTGTTTGTATC 240  
QY 241 TTGTTGACTATGAGAAATCTAGAAATCTAGAAATCTAGAAATCTAGAAATCTAGAAAT 300  
DB 241 TTGTTGACTATGAGAAATCTAGAAATCTAGAAATCTAGAAATCTAGAAATCTAGAAAT 300  
QY 301 ATTTTGTATGTGATGATGCCAAGTTTATAGGACTTAAAGAGAGATCTACTTCAGAGC 360  
DB 301 ATTTTGTATGTGATGATGCCAAGTTTATAGGACTTAAAGAGAGATCTACTTCAGAGC 360  
QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGCAGGAGCAGGTTCTGTGG 420  
DB 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGCAGGAGCAGGTTCTGTGG 420  
QY 421 AGGCAGAACCCAGAAATCTGAAGATGAAGCAAAAGCAAAATTCAGTCCCTTCTCCATG 480  
DB 421 AGGCAGAACCCAGAAATCTGAAGATGAAGCAAAAGCAAAATTCAGTCCCTTCTCCATG 480  
QY 481 AAATGGTACACCGAAGCATGTTGAGGAGAGAGACTTGCAACAGAAAGATGAGCCACAG 540  
DB 481 AAATGGTACACCGAAGCATGTTGAGGAGAGAGACTTGCAACAGAAAGATGAGCCACAG 540  
QY 541 GAGAACCAACAAGAGGATGATGAGTTTCTTATGCGCATCTGATGTAGATGATGATTTG 600  
DB 541 GAGAACCAACAAGAGGATGATGAGTTTCTTATGCGCATCTGATGTAGATGATGATTTG 600  
QY 601 AGACCTGGAACTGAGATCTCATGAAGAACCGCATAGTTTACACGTTGGAGAGA 660  
DB 601 AGACCTGGAACTGAGATCTCATGAAGAACCGCATAGTTTACACGTTGGAGAGA 660  
QY 661 CAGTTTCAACAAGACTGTAAATCAGATATGGAAGAGATGATGTCTGAGCAGGAAATCCAG 720  
DB 661 CAGTTTCAACAAGACTGTAAATCAGATATGGAAGAGATGATGTCTGAGCAGGAAATCCAG 720

QY 721 ATTCCAGTGAACACAGTAGTAGAAGATGAAGATTGACCATCATACAGATGATTAACAT 780  
DB 721 ATTCCAGTGAACACAGTAGTAGAAGATGAAGATTGACCATCATACAGATGATTAACAT 780  
QY 781 ACCAGTCTTATGAGGAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA 840  
DB 781 ACCAGTCTTATGAGGAACAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCA 840  
QY 841 CAGAACTAACTGCTCCCTCCCTGAGGATAATCTGTAGAAATTCACAGGTAATTTGTAGAAG 900  
DB 841 CAGAACTAACTGCTCCCTCCCTGAGGATAATCTGTAGAAATTCACAGGTAATTTGTAGAAG 900  
QY 901 AAGTAAGCAATTTTCTGTGGAAGAACACAGCAGGAAGTACACACAGAAACAAATAGAAAA 960  
DB 901 AAGTAAGCAATTTTCTGTGGAAGAACACAGCAGGAAGTACACACAGAAACAAATAGAAAA 960  
QY 961 CAGATGATCCAGAAACAAAGCAAAAGTTTAAAGAAAAAGAGCTAAATCTTTTAAATAAT 1020  
DB 961 CAGATGATCCAGAAACAAAGCAAAAGTTTAAAGAAAAAGAGCTAAATCTTTTAAATAAT 1020  
QY 1021 TTGATAAGACTATTTAAAGCTGAACTTTGATGCTGCAGAAAAAATCTCCGTAAAGGGGAAAA 1080  
DB 1021 TTGATAAGACTATTTAAAGCTGAACTTTGATGCTGCAGAAAAAATCTCCGTAAAGGGGAAAA 1080  
QY 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGAAATACCTCTCAGAGTCCAGAG 1140  
DB 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGAAATACCTCTCAGAGTCCAGAG 1140  
QY 1141 CAAGATATGGAAGGGCGCAGTGGAGATGATTTGGCTGAGAAAGAGAGAGAAATTAAGG 1200  
DB 1141 CAAGATATGGAAGGGCGCAGTGGAGATGATTTGGCTGAGAAAGAGAGAGAAATTAAGG 1200  
QY 1201 TGCTACGTGAGGCATCGAGACCTTCAAGAGGTGGCGAGCCTACCTGATGTCCTGCGAG 1260  
DB 1201 TGCTACGTGAGGCATCGAGACCTTCAAGAGGTGGCGAGCCTACCTGATGTCCTGCGAG 1260  
QY 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAAACTTCTAGGTTCATATGA 1320  
DB 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAAACTTCTAGGTTCATATGA 1320  
QY 1321 GAGGTTCCCTGCTTACCTTCAGAGATTTAGTTCAACTATTTCCCAATGATCTTCTTAA 1380  
DB 1321 GAGGTTCCCTGCTTACCTTCAGAGATTTAGTTCAACTATTTCCCAATGATCTTCTTAA 1380  
QY 1381 AAAATGACCTTGGCGTGGGATACCTCTTGTAGAGATAATGCAATTCGCAAGAAAGTTT 1440  
DB 1381 AAAATGACCTTGGCGTGGGATACCTCTTGTAGAGATAATGCAATTCGCAAGAAAGTTT 1440  
QY 1441 ATGAAGAGGTGCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTTATGGCTTCA 1500  
DB 1441 ATGAAGAGGTGCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTTATGGCTTCA 1500  
QY 1501 TCCTGAAAGGCACAGAACAAATTTGCTGAGAGCATCCCATATTTAAAGAGAAATAGAAAT 1560  
DB 1501 TCCTGAAAGGCACAGAACAAATTTGCTGAGAGCATCCCATATTTAAAGAGAAATAGAAAT 1560  
QY 1561 CCGGAGATCTTGGCATGTGATGGGATTTTATTTCCACTGGGGATGCCATGCGAGA 1620  
DB 1561 CCGGAGATCTTGGCATGTGATGGGATTTTATTTCCACTGGGGATGCCATGCGAGA 1620  
QY 1621 GGGTTGGGAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCACAAAGAGAGACATTTTG 1680  
DB 1621 GGGTTGGGAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCACAAAGAGAGACATTTTG 1680  
QY 1681 CATCTGTCTGCAACGCTCACTCTACATGATGATGATGATGATGATGATGATGATGATG 1740  
DB 1681 CATCTGTCTGCAACGCTCACTCTACATGATGATGATGATGATGATGATGATGATGATG 1740  
QY 1741 CCCCAAAAGAAACGGGCTTACACAGATTTAGTAAAGTCTTTTAAAGAAAGAACTGGAAGTTAA 1800  
DB 1741 CCCCAAAAGAAACGGGCTTACACAGATTTAGTAAAGTCTTTTAAAGAAAGAACTGGAAGTTAA 1800  
QY 1801 TCCGAGATGAAGGCCCTTTCAGTGTGATGATGAAGCAAGAGGCTCTCTTCTGCTGAGGATG 1860  
DB 1801 TCCGAGATGAAGGCCCTTTCAGTGTGATGATGAAGCAAGAGGCTCTCTTCTGCTGAGGATG 1860

DB 1801 TCCGAGATGAAGGCTTCGATGATGAAGCCAAAGGTCCTCTCTCCCTGAGGATG 1860  
QY 1861 AAAACCTGAGGAAAAGGGGACTGGAGCCAGTTCACGCTGTGCGACGAAGAGAA 1920  
DB 1861 AAAACCTGAGGAAAAGGGGACTGGAGCCAGTTCACGCTGTGCGACGAAGAGAA 1920  
QY 1921 ATGAAATGCTGCAAGAGGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGAGA 1980  
DB 1921 ATGAAATGCTGCAAGAGGAGCTCTTAAACCTGTACCTTACTAGAAAAGTTCCCGAGA 1980  
QY 1981 CAACAGGATGAGAGAGGACAGATCAATATTTCCATCATGACCCCGGAGCTCACGTGT 2040  
DB 1981 CAACAGGATGAGAGAGGACAGATCAATATTTCCATCATGACCCCGGAGCTCACGTGT 2040  
QY 2041 GGCGGACACAGGGGCCCCAACAATGCGAGGCTCCGAATGACCTGGGCTGGTATCCCA 2100  
DB 2041 GGCGGACACAGGGGCCCCAACAATGCGAGGCTCCGAATGACCTGGGCTGGTATCCCA 2100  
QY 2101 AGGAAGCTGCAAGATTCGATGTGCCACAGAGACAGGACCTGGGAGGAAGCAAGGTGC 2160  
DB 2101 AGGAAGCTGCAAGATTCGATGTGCCACAGAGACAGGACCTGGGAGGAAGCAAGGTGC 2160  
QY 2161 TCATCTTTGATGACTCCTTTTGAGCAGAGGTATGCGAGGATGCTCATCTTTCCGGCTGA 2220  
DB 2161 TCATCTTTGATGACTCCTTTTGAGCAGAGGTATGCGAGGATGCTCATCTTTCCGGCTGA 2220  
QY 2221 TATTTCATGCTGGATGTGGCATTCGGAACTGACACACAGCAGAGAGCGCTTCACAG 2280  
DB 2221 TATTTCATGCTGGATGTGGCATTCGGAACTGACACACAGCAGAGAGCGCTTCACAG 2280  
QY 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAAAACCTCTGGAGAGA 2324  
DB 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAAAACCTCTGGAGAGA 2324  
RESULT 3  
ID ADA00640 standard; cDNA; 2324 BP.  
AC ADA00640;  
XX  
DT 06-NOV-2003 (first entry)  
DE cDNA encoding human aspartyl (asparaginyl) beta-hydroxylase (HAAH).  
XX Tumour growth inhibition; human aspartyl (asparaginyl) beta-hydroxylase;  
KW HAAH hydroxylation; NOTCH polypeptide;  
KW epidermal growth factor-like repeat; EGF-like repeat; tumour cell;  
KW malignant neoplasm; colon cancer; breast cancer; pancreatic cancer;  
KW liver cancer; cancer of the bile duct; cancer the central nervous system;  
KW CNS; cytostatic; gene; ss; human.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 12..2288  
FT /\*tag= a  
FT /product= "HAAH"  
XX  
PN US2003031670-A1.  
XX  
PD 13-FEB-2003.  
XX  
PF 08-NOV-1999; 99US-00436184.  
XX  
PR 08-NOV-1999; 99US-00436184.  
XX (WAND/) WANDS J. R.  
PA (DMON/) DE LA MONTE S. M.  
PA (INCE/) INCE N.  
PA (CARL/) CARLSON R. I.

PI Wands JR, De La Monte SM, Ince N, Carlson RI;  
XX WPI; 2003-605701/57.  
DR P-PSDB; ADA00639.  
XX  
XX Inhibiting tumor growth or killing tumor cells (e.g. cancer of the colon,  
PT breast, pancreatic, liver or the central nervous system), by  
PT administering an inhibitor of the human aspartyl (asparaginyl) beta-  
PT hydroxylase.  
XX  
PS Disclosure; Page 3-4; 30pp; English.  
XX  
CC The present invention relates to a method for inhibiting tumour growth in  
CC a mammal. The method comprises administering to the mammal a compound,  
CC which inhibits the expression or enzymatic activity of a human aspartyl  
CC (asparaginyl) beta-hydroxylase (HAAH). The compound may inhibit HAAH  
CC hydroxylation of a NOTCH polypeptide. In particular, the compound may  
CC inhibit hydroxylation of an epidermal growth factor (EGF)-like repeat  
CC sequence in a NOTCH polypeptide. The methods are useful for inhibiting  
CC tumour growth or killing tumour cells, or for diagnosing or  
CC prognosticating a malignant neoplasm. In particular, the tumour or  
CC neoplasm is colon cancer, breast cancer, pancreatic cancer, liver cancer,  
CC cancer of the bile ducts, or cancer or tumour of the central nervous  
CC system (CNS). The present sequence encodes HAAH.  
XX  
SQ Sequence 2324 BP; 753 A; 463 C; 625 G; 483 T; 0 U; 0 Other;  
Query Match 100.0%; Score 2324; DB 9; Length 2324;  
Best Local Similarity 100.0%; Pred No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGACCGTGCATATGGCCCGAGCGTAGAATGCCAAGAGCAGCGGCAACAGCAGCAGCGG 60  
DB 1 CGGACCGTGCATATGGCCCGAGCGTAGAATGCCAAGAGCAGCGGCAACAGCAGCAGCGG 60  
QY 61 GCTCCGGCAGCGGTAGCACAGAGTGGCGGCGAGCAGCGCCCGGGCCCGGAGAGACAA 120  
DB 61 GCTCCGGCAGCGGTAGCACAGAGTGGCGGCGAGCAGCGCCCGGGCCCGGAGAGACAA 120  
QY 121 AGCATGGAGGACACAAGAAATGGAGAAAGCGGACTCTCGGAACTTCATTCTTCACGT 180  
DB 121 AGCATGGAGGACACAAGAAATGGAGAAAGCGGACTCTCGGAACTTCATTCTTCACGT 180  
QY 181 GGTATTGCTGATGTCATTTGCTGGCGCTGAGACTCTGAGTCTGCTGTTGTTGATC 240  
DB 181 GGTATTGCTGATGTCATTTGCTGGCGCTGAGACTCTGAGTCTGCTGTTGTTGATC 240  
QY 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGGTGGAG 300  
DB 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGGTGGAG 300  
QY 301 ATTTTGTATGATGATGTCACAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTCAGAGC 360  
DB 301 ATTTTGTATGATGATGTCACAAAGTTTATTAGGACTTTAAAGAGAGATCTACTTCAGAGC 360  
QY 361 CAGCAGTCCCGCCAGAGAGAGGCTGAGCCACACACTGAGCCCGAGGACAGGTTCTCTGAG 420  
DB 361 CAGCAGTCCCGCCAGAGAGAGGCTGAGCCACACACTGAGCCCGAGGACAGGTTCTCTGAG 420  
QY 421 AGGCAGAACCCGAGAAATATCGAAGTGAAGCAAGAAACAAATTCAGTCCCTCTCCATG 480  
DB 421 AGGCAGAACCCGAGAAATATCGAAGTGAAGCAAGAAACAAATTCAGTCCCTCTCCATG 480  
QY 481 AAATGGTACACGCAAGAACATGTTGAGGGGAGAGACTTGCACAAAGAAAGATGGACCCACAG 540  
DB 481 AAATGGTACACGCAAGAACATGTTGAGGGGAGAGACTTGCACAAAGAAAGATGGACCCACAG 540  
QY 541 GAGAACCAACCAAGAGGATGATGATTTCTTATGGGACTGATGATGATGATGATGATG 600  
DB 541 GAGAACCAACCAAGAGGATGATGATTTCTTATGGGACTGATGATGATGATGATGATG 600  
QY 601 AGACCTGGAACCTGAAGTATCTCATGAGAAACCGAGCATAGTACACCTGGAGAGA 660  
DB 601 AGACCTGGAACCTGAAGTATCTCATGAGAAACCGAGCATAGTACACCTGGAGAGA 660

```
Db 601 AGACCCGTGGAACCTGAAGTATCTCATGAAGAAACCGAGCATAGTTTACACGTGGAAGAGA 660
Qy 661 CAGTTTCACAGACTGTAAATCAGGATATGGAAGAGATGATGTCCTGAGCAGGAAATCCAG 720
Db 661 CAGTTTCACAGACTGTAAATCAGGATATGGAAGAGATGATGTCCTGAGCAGGAAATCCAG 720
Qy 721 ATCCAGTGAACCAAGTAGTAGAAGATGAAGATGCAACCATGATACAGATGATTAACAT 780
Db 721 ATCCAGTGAACCAAGTAGTAGAAGATGAAGATGCAACCATGATACAGATGATTAACAT 780
Qy 781 ACCAAGTCTATGAGGAACCAAGCAGTATATGAACCTCTAGAAAATGAAGGGATAGAAATCA 840
Db 781 ACCAAGTCTATGAGGAACCAAGCAGTATATGAACCTCTAGAAAATGAAGGGATAGAAATCA 840
Qy 841 CAGAAGTAACCTGCTCCCTCAGGATATATCTGTAGAAAGATTACAGGTAATTTCTAGAAG 900
Db 841 CAGAAGTAACCTGCTCCCTCAGGATATATCTGTAGAAAGATTACAGGTAATTTCTAGAAG 900
Qy 901 AAGTAAGCATTTTCTGTGGAAGAACAGCAGGAAGTACCACAGAAACAAATAGAAAAA 960
Db 901 AAGTAAGCATTTTCTGTGGAAGAACAGCAGGAAGTACCACAGAAACAAATAGAAAAA 960
Qy 961 CAGATGATCCAGAACAAAGCAAAAGTTAAGAAAAGAGCCTTAAACTTTTAAATAAT 1020
Db 961 CAGATGATCCAGAACAAAGCAAAAGTTAAGAAAAGAGCCTTAAACTTTTAAATAAT 1020
Qy 1021 TTGATAAGACTATTAAAGCTGAACTTGATGCTGCAGAAAAAACTCCGTAAAAAGGGAAAAA 1080
Db 1021 TTGATAAGACTATTAAAGCTGAACTTGATGCTGCAGAAAAAACTCCGTAAAAAGGGAAAAA 1080
Qy 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGCAAAATACCCCTCAGAGTCCAGAG 1140
Db 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGCAAAATACCCCTCAGAGTCCAGAG 1140
Qy 1141 CAAGATATGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAAAGAGAGATTAATGAGG 1200
Db 1141 CAAGATATGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAAAGAGAGATTAATGAGG 1200
Qy 1201 TGCTACGTGGAGCCATCCAGACCTACCAAGAGGTGGCCAGCCTACCTGATGCTCCCTGCAG 1260
Db 1201 TGCTACGTGGAGCCATCCAGACCTACCAAGAGGTGGCCAGCCTACCTGATGCTCCCTGCAG 1260
Qy 1261 ACCTGCTGAAGCTGAGTTGAAGCGTGCCTCAGACAGGCAACAATTTCTAGGTCAATATGA 1320
Db 1261 ACCTGCTGAAGCTGAGTTGAAGCGTGCCTCAGACAGGCAACAATTTCTAGGTCAATATGA 1320
Qy 1321 GAGGTCCTCCTCTACCTGCGAGGATTTAGTTCACATATTTCCCAATGATACTTCCCTTAA 1380
Db 1321 GAGGTCCTCCTCTACCTGCGAGGATTTAGTTCACATATTTCCCAATGATACTTCCCTTAA 1380
Qy 1381 AAAATGACCTTGGCGTGGGATACCTCTTGATAGAGATAATGACAATGCAAAAGAAAGTTT 1440
Db 1381 AAAATGACCTTGGCGTGGGATACCTCTTGATAGAGATAATGACAATGCAAAAGAAAGTTT 1440
Qy 1441 ATGAAGAGGTGCTGAGTGTGACACCTAATATGAGCTTTGCTTAAAGTCAATATGCTTCA 1500
Db 1441 ATGAAGAGGTGCTGAGTGTGACACCTAATATGAGCTTTGCTTAAAGTCAATATGCTTCA 1500
Qy 1501 TCCTGAAGGCACAGAACAAATTTGCTGAGAGCATCCCATATTTTAAAGGAAGGAATAGAAT 1560
Db 1501 TCCTGAAGGCACAGAACAAATTTGCTGAGAGCATCCCATATTTTAAAGGAAGGAATAGAAT 1560
Qy 1561 CCGGAGATCTCTGGCACTGATGATGGAGATTTTATTTCCACCTGGGGATGCCATGCAGA 1620
Db 1561 CCGGAGATCTCTGGCACTGATGATGGAGATTTTATTTCCACCTGGGGATGCCATGCAGA 1620
Qy 1621 GGGTTGGGAAACAAAGAGGCATATAAGTGTATGAGCTTTGGGCACAAAGAGAGACACTTTG 1680
Db 1621 GGGTTGGGAAACAAAGAGGCATATAAGTGTATGAGCTTTGGGCACAAAGAGAGACACTTTG 1680
Qy 1681 CATCTGCTGCAACGCTCACTCTACAAATGTAATGGAATGGAATGGAATGGAATGGAATGGA 1740
Db 1681 CATCTGCTGCAACGCTCACTCTACAAATGTAATGGAATGGAATGGAATGGAATGGAATGGA
```

```
Qy 1741 CCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAAGAAACTGGAAGTTAA 1800
Db 1741 CCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAAGAAACTGGAAGTTAA 1800
Qy 1801 TCCGAGATGAAGGCTTTGCAAGTATGATGAATGAAGCAAGGCTCTTCTCCTCGCTGAGATG 1860
Db 1801 TCCGAGATGAAGGCTTTGCAAGTATGATGAATGAAGCAAGGCTCTTCTCCTCGCTGAGATG 1860
Qy 1861 AAAACCTGAGGGAAGGAGGAGCTGAGCCAGTTCACGCTGTGCGACGCAAGGAAGAAAGAA 1920
Db 1861 AAAACCTGAGGGAAGGAGGAGCTGAGCCAGTTCACGCTGTGCGACGCAAGGAAGAAAGAA 1920
Qy 1921 ATGAAAAATGCTGCAAAAGGAGCTCTTAAAACTGTACCTTACTAGAAAAAGTTCCCGGAGA 1980
Db 1921 ATGAAAAATGCTGCAAAAGGAGCTCTTAAAACTGTACCTTACTAGAAAAAGTTCCCGGAGA 1980
Qy 1981 CAACAGAGATGCAAGAGGAGCAGATCAAAATATTTCCATCATGCAACCCCGGAGCTCACGTGT 2040
Db 1981 CAACAGAGATGCAAGAGGAGCAGATCAAAATATTTCCATCATGCAACCCCGGAGCTCACGTGT 2040
Qy 2041 GGCCGCAACACAGGCCCCACAAAACCTGCAGGCTCCGAATGCACCTGGGCTTGGTGATTTCCCA 2100
Db 2041 GGCCGCAACACAGGCCCCACAAAACCTGCAGGCTCCGAATGCACCTGGGCTTGGTGATTTCCCA 2100
Qy 2101 AGGAAGGCTGCAAGATTCGATGTCGCAACAGAGACCAAGACCTTGGGAGAGAGGCAAGGTGC 2160
Db 2101 AGGAAGGCTGCAAGATTCGATGTCGCAACAGAGACCAAGACCTTGGGAGAGAGGCAAGGTGC 2160
Qy 2161 TCATCTTTGATGATCTCTTTGAGCAGCAGGATGTCGAGGATGTCGAGGATGTCGAGGATGTC 2220
Db 2161 TCATCTTTGATGATCTCTTTGAGCAGCAGGATGTCGAGGATGTCGAGGATGTCGAGGATGTC 2220
Qy 2221 TATTATCGTGGATGTCGTCATCCGGAACCTGACACACAGCAGACGAGCAGCCTTCCAG 2280
Db 2221 TATTATCGTGGATGTCGTCATCCGGAACCTGACACACAGCAGACGAGCAGCCTTCCAG 2280
Qy 2281 CAATTTAGCATGAATTCATGCAAGCTTTGGGAAACTCTGGAGAGA 2324
Db 2281 CAATTTAGCATGAATTCATGCAAGCTTTGGGAAACTCTGGAGAGA 2324
```

## RESULT 4

ADP23914  
ID ADP23914 standard; cDNA; 2452 BP.

AC ADP23914;

XX DT 18-NOV-2004 (first entry)

XX DE PRO polypeptide encoding cDNA SEQ ID NO:1092.

XX ss: gene; PRO; antinflammatory; antirheumatic; antirheumatic;  
XX immunosuppressive; osteopathic; antidiabetic; dermatological;  
XX antipsoriatic; antiallergic; antisthmatic; hepatotropic; respiratory;  
XX gene therapy; immune system.

OS Unidentified.

PN WO2004041170-A2.

XX 21-MAY-2004.

XX 30-OCT-2003; 2003WO-US034312.

XX 01-NOV-2002; 2002US-0423394P.

XX (GETH ) GENENTECH INC.

XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;  
XX Wu TD;

XX WPI; 2004-419628/39.

DR P-PSDB; ADP23915.  
XX New PRO polypeptides and polynucleotides, useful for treating e.g.  
PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated  
PT renal disease, or demyelinating diseases of the central or peripheral  
PT nervous system.  
XX Claim 1; SEQ ID NO 1092; 2940pp; English.  
XX The invention relates to a novel isolated nucleic acid and the PRO  
CC polypeptide encoded by it. A protein of the invention has  
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,  
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,  
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide  
CC of the invention may have a use in gene therapy. The PRO polypeptide, its  
CC agonist, antagonist, or antibody that specifically binds to the  
CC polypeptide is useful for treating an immune related disorder such as  
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,  
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an  
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic  
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune  
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal  
CC disease, a demyelinating disease of the central or peripheral nervous  
CC system, idiopathic demyelinating polynuropathy, Guillain-Barre syndrome,  
CC a chronic inflammatory demyelinating polynuropathy, a hepatobiliary  
CC disease, infectious or autoimmune chronic active hepatitis, primary  
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,  
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's  
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin  
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic  
CC disease, asthma, allergic rhinitis, atopic dermatitis, food  
CC hypersensitivity, urticaria, an immunologic disease of the lung,  
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity  
CC pneumonitis, a transplantation associated disease, graft rejection or  
CC Graft-versus-host disease. The present sequence encodes a PRO protein of  
CC the invention.  
XX  
XX Sequence 2452 BP; 779 A; 507 C; 657 G; 509 T; 0 U; 0 Other;  
Query Match 95.6%; Score 2222; DB 13; Length 2452;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2322; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CGGACCGTGAATGGCCCGACGCTAAGATGCCAAGCAGCAGCGGCAACAGCAGCAGCG 60  
DB 67 CGGACCGTGAATGGCCCGACGCTAAGATGCCAAGCAGCAGCGGCAACAGCAGCAGCG 126  
QY 61 GCTCCGCGCGGTAGCAGAGTGGCGGCGCAGCAGCAGCGCCCGGCGCCGAGAGACAA 120  
DB 127 GCTCCGCGCGGTAGCAGAGTGGCGGCGCAGCAGCAGCGCCCGGCGCCGAGAGACAA 186  
QY 121 AGCATGGAGGACACAAAGATGGAGGAAAGCGGCACTCTCGGGAACCTTCATCTTCACGT 180  
DB 187 AGCATGGAGGACACAAAGATGGAGGAAAGCGGCACTCTCGGGAACCTTCATCTTCACGT 246  
QY 181 GGTATTATGGTATGCTGCGGCGCTCGACACTGTAGCTGTGCTTGGTTTGGTTC 240  
DB 247 GGTATTATGGTATGCTGCGGCGCTCGACACTGTAGCTGTGCTTGGTTTGGTTC 306  
QY 241 TTGTTGACTATAGGAGTCTTAGGAAACTAGGAATCTAGTCTGTGCTGTGATGGAG 300  
DB 307 TTGTTGACTATAGGAGTCTTAGGAAACTAGGAATCTAGTCTGTGCTGTGATGGAG 366  
QY 301 ATTTTGTATGGATGATGCAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 360  
DB 367 ATTTTGTATGGATGATGCAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 426  
QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGAGGTTCTGTGG 420  
DB 427 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGAGAGGTTCTGTGG 486  
QY 421 AGGCAGAACCCAGAAATATCGAAGATGAAGCAAGAAACAAATTCAGTCCCTTCATG 480

DB 487 AGGCAGAACCCAGAAATATCGAAGATGAAGCAAGAAACAAATTCAGTCCCTTCATG 546  
QY 481 AAATGGTACCGCAGAACATGTTTGGGGGAGAGACTTGGCAACAAGAGATGGACCCACAG 540  
DB 547 AAATGGTACCGCAGAACATGTTTGGGGGAGAGACTTGGCAACAAGAGATGGACCCACAG 606  
QY 541 GAGAACCAACAAAGAGAGATGATGATTTCTTTATGGCGACTGATGATGATGATGATTTG 600  
DB 607 GAGAACCAACAAAGAGAGATGATGATTTCTTTATGGCGACTGATGATGATGATGATTTG 666  
QY 601 AGACCCCTCGAACCTGGAAGTATCTCATGAAGAAACCGAGCATAGTTACCAGCTGGAAGAGA 660  
DB 667 AGACCCCTCGAACCTGGAAGTATCTCATGAAGAAACCGAGCATAGTTTCCAGCTGGAAGAGA 726  
QY 661 CAGTTTTCACAGACTGTAAATCAGGATATCGAAGAGATGATGCTGAGCAGGAAATTCAG 720  
DB 727 CAGTTTTCACAGACTGTAAATCAGGATATCGAAGAGATGATGCTGAGCAGGAAATTCAG 786  
QY 721 ATTCCAGTGAACCAAGTATGAGAAGATGAAGATTCACCATGATACAGATGATGATTAACAT 780  
DB 787 ATTCCAGTGAACCAAGTATGAGAAGATGAAGATTCACCATGATACAGATGATGATTAACAT 846  
QY 781 ACCAAGTCTATGAGGAAACGAGCATATATGAACCTCTAGAAAATGAAGGGATAGAATCA 840  
DB 847 ACCAAGTCTATGAGGAAACGAGCATATATGAACCTCTAGAAAATGAAGGGATAGAATCA 906  
QY 841 CAGAAGTAACCTGCTCCCTCGAGGATTAATCTGTAGAGATTCACAGTAAATTTGTAGAAG 900  
DB 907 CAGAAGTAACCTGCTCCCTCGAGGATTAATCTGTAGAGATTCACAGTAAATTTGTAGAAG 966  
QY 901 AAGTAAGCATTTTCTGTGTGGAAGAACAGCAGGAAAGTACCAACAGAAAACAAATAGAAAAA 960  
DB 967 AAGTAAGCATTTTCTGTGTGGAAGAACAGCAGGAAAGTACCAACAGAAAACAAATAGAAAAA 1026  
QY 961 CAGATGATCCAGAACAAAAAGCAAAGGTTAAGAAAAAGAGCCCTTAACTTTTAAATAAAT 1020  
DB 1027 CAGATGATCCAGAACAAAAAGCAAAGGTTAAGAAAAAGAGCCCTTAACTTTTAAATAAAT 1086  
QY 1021 TTGATAAGACTATTAAAGCTGAACCTTGATGCTGCAGAAAAAATCCGTAAAAAGGGAAAAA 1080  
DB 1087 TTGATAAGACTATTAAAGCTGAACCTTGATGCTGCAGAAAAAATCCGTAAAAAGGGAAAAA 1146  
QY 1081 TTGAGGAAGCAGTGAATGATTAAGCACTAGTACGCAATACCTCAGAGTCCACAGAG 1140  
DB 1147 TTGAGGAAGCAGTGAATGATTAAGCACTAGTACGCAATACCTCAGAGTCCACAGAG 1206  
QY 1141 CAGATATGGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGAAAGTAAATGAGG 1200  
DB 1207 CAGATATGGGAAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGAAAGTAAATGAGG 1266  
QY 1201 TGCTACCTGAGAGCCATCGAGAGCTACCAAGAGAGTGGCGAGCCTACCTGATGCTCCCTGCGAG 1260  
DB 1267 TGCTACCTGAGAGCCATCGAGAGCTACCAAGAGAGTGGCGAGCCTACCTGATGCTCCCTGCGAG 1326  
QY 1261 ACCTGCTGAAGCTGAGTTTGAAGCCTCGCTCAGACAGCAACAAATTTCTAGGTCATATGA 1320  
DB 1327 ACCTGCTGAAGCTGAGTTTGAAGCCTCGCTCAGACAGCAACAAATTTCTAGGTCATATGA 1386  
QY 1321 CAGGTTCCCTGCTTACCTTCAGAGATTAAGTCAATTAATTTCCCAATGATATCTTCCTTAA 1380  
DB 1387 CAGGTTCCCTGCTTACCTTCAGAGATTAAGTCAATTAATTTCCCAATGATATCTTCCTTAA 1446  
QY 1381 AAAATGACCTTGGCGTGGGATACCTCTTGTATAGAGATAATGACAAATGCAAGAAAGTTT 1440  
DB 1447 AAAATGACCTTGGCGTGGGATACCTCTTGTATAGAGATAATGACAAATGCAAGAAAGTTT 1506  
QY 1441 ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGCTTGTAAAGTCCCAATTTAGGCTTCA 1500  
DB 1507 ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGCTTGTAAAGTCCCAATTTAGGCTTCA 1566  
QY 1501 TCCTGAGGCGACAGAACAAATTTGCTGAGAGCATCCCATATTTTAAAGAGGAATAGAT 1560  
DB 1567 TCCTGAGGCGACAGAACAAATTTGCTGAGAGCATCCCATATTTTAAAGAGGAATAGAT 1626

QY 1561 CCGGAGATCTCGCACTGATGATGGGATTTTATTTCCACCTGGGGATGCCATGCGA 1620  
Db 1627 CCGGAGATCTCGCACTGATGATGGGATTTTATTTCCACCTGGGGATGCCATGCGA 1686  
QY 1621 GGGTTGGGAACAAGAGGCAATATAAGTGGTATAGCTTGGGCAACAAGAGGACACTTTG 1680  
Db 1687 GGGTTGGGAACAAGAGGCAATATAAGTGGTATAGCTTGGGCAACAAGAGGACACTTTG 1746  
QY 1681 CATCTGTGCAAGCGCTCACTTACATGTGATGATGACATGAAGACACAGCTTGGTGA 1740  
Db 1747 CATCTGTGCAAGCGCTCACTTACATGTGATGATGACATGAAGACACAGCTTGGTGA 1806  
QY 1741 CCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTTGAAGTTAA 1800  
Db 1807 CCCAAAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTTGAAGTTAA 1866  
QY 1801 TCCGAGATGAAGGCTTGCATGTATGATGATGAAGCAAGAGTCTCTTCTGCTGAGATG 1860  
Db 1867 TCCGAGATGAAGGCTTGCATGTATGATGATGAAGCAAGAGTCTCTTCTGCTGAGATG 1926  
QY 1861 AAAACCTGAGGAAAGGGGACTGGAGCCAGTTTCAGCTGTGCAGCAAGCAAGAAAGAA 1920  
Db 1927 AAAACCTGAGGAAAGGGGACTGGAGCCAGTTTCAGCTGTGCAGCAAGCAAGAAAGAA 1986  
QY 1921 ATGAAATGCTGCAAGGAGCTCTTAAACCTGTACCTTACTAGAAAGTTTCCCGAGA 1980  
Db 1987 ATGAAATGCTGCAAGGAGCTCTTAAACCTGTACCTTACTAGAAAGTTTCCCGAGA 2046  
QY 1981 CAACAGGATGAGAAAGAGAGAGATCAAAATATTCATCATGACCCCGGAGTCACTGT 2040  
Db 2047 CAACAGGATGAGAAAGAGAGAGATCAAAATATTCATCATGACCCCGGAGTCACTGT 2106  
QY 2041 GGCCGACACAGGGGCCCCAAGCTGCGAGTCCGAGTCCGAGTCTGGCTTGGTATCCCA 2100  
Db 2107 GGCCGACACAGGGGCCCCAAGCTGCGAGTCCGAGTCCGAGTCTGGCTTGGTATCCCA 2166  
QY 2101 AGGAAGGCTGCAAGATTCATGTGTCACAGAGACAGGACCTGGGAGAGGCAAGGTGC 2160  
Db 2167 AGGAAGGCTGCAAGATTCATGTGTCACAGAGACAGGACCTGGGAGAGGCAAGGTGC 2226  
QY 2161 TCATCTTTGATGACTCTTTGAGCAGAGGATGTCAGAGTGCCTCATCTTTCGGCTGA 2220  
Db 2227 TCATCTTTGATGACTCTTTGAGCAGAGGATGTCAGAGTGCCTCATCTTTCGGCTGA 2286  
QY 2221 TATTCATCTGTGATCTGTGGCATCCGAACTGACACACAGCAGAGACGAGCCTTCCAG 2280  
Db 2287 TATTCATCTGTGATCTGTGGCATCCGAACTGACACACAGCAGAGACGAGCCTTCCAG 2346  
QY 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTTCTGGAGAGA 2324  
Db 2347 CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTTCTGGAGAGA 2390

RESULT 5  
ID ADR97347  
XX ADR97347 standard; DNA; 2452 BP.  
AC ADR97347;  
XX ADR97347;  
DT 02-DEC-2004 (first entry)  
DE Human ASPH DNA, an apoptosis related target Seq 55.  
XX  
KW gene; ds; human; apoptosis; cancer; inflammation; autoimmune;  
KW neurodegenerative disorder; cytostatic; antineoplastic;  
KW immunosuppressive; neuroprotective; gene therapy; ASPH;  
KW aspartate beta hydroxylase.  
XX  
OS Homo sapiens.  
XX  
PN WO2004078783-A2.  
XX

PD 16-SEP-2004.  
XX  
PF 05-MAR-2004; 2004WO-GB000957.  
XX  
PR 07-MAR-2003; 2003GB-00005267.  
XX  
PA (EIRX-) EIRX THERAPEUTICS LTD.  
XX  
PI Murphy FU, Sheehan DE, Keating KE, Hayes I, Seera L;  
XX  
DR WPI: 2004-662402/64.  
XX  
DR P-PSDB; ADR97348.  
XX  
PT Identifying an agent that modulates the function of an apoptosis-  
PT associated polypeptide, useful for diagnosing or treating e.g. cancer,  
PT comprises comparing the binding of the polypeptide to the candidate agent  
PT and to a control agent.  
XX  
PS Claim 2; SEQ ID NO 55; 304pp; English.  
XX  
CC This invention relates to novel agents that modulates the function of  
CC human apoptosis-associated proteins specified within the specification.  
CC Specifically, it refers to a method for the identification of target  
CC genes whose expression is correlated with an early stage in the  
CC regulation of apoptosis. The present invention describes a method of  
CC contacting either candidate agents or control agents to the target genes  
CC and assessing the difference of binding and inhibitory activity, where  
CC the candidate agent is selected from a low molecular weight organic  
CC molecule, an antibody or fragment thereof, an antisense oligonucleotide,  
CC a small inhibitory dsRNA, or a ribozyme. As such, the compositions and  
CC methods are useful for diagnosing and treating diseases or conditions  
CC associated with abnormal apoptosis in mammalian tissue, such as cancer,  
CC inflammation, autoimmune or neurodegenerative disorders. Accordingly,  
CC they exhibit cytostatic, antineoplastic, immunosuppressive and  
CC neuroprotective activities. These may also be used for drug screening  
CC purposes and in gene therapy. This polynucleotide sequence is a human  
CC target gene associated with the regulation of apoptosis that can be  
CC modulated by novel agents of the invention.  
XX  
SQ Sequence 2452 BP; 779 A; 507 C; 657 G; 509 T; 0 U; 0 Other;

Query Match 95.6%; Score 2222; DB 13; Length 2452;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2322; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 CGGACCGTGCATCGCCAGCGTAAAGATGCCAAGAGCAGCGCAACAGCAGCAGCG 60  
Db 67 CGGACCGTGCATCGCCAGCGTAAAGATGCCAAGAGCAGCGCAACAGCAGCAGCG 126  
QY 61 GCTCCGGCAGCGGTAGCACGAGTCCGGGAGGAAAGCGGACCTCTCGGAACTTCACTTTCACGT 120  
Db 127 GCTCCGGCAGCGGTAGCACGAGTCCGGGAGGAAAGCGGACCTCTCGGAACTTCACTTTCACGT 186  
QY 121 AGCATGGAGGACACAGAAATGGGAGGAAAGCGGACCTCTCGGAACTTCACTTTCACGT 180  
Db 187 AGCATGGAGGACACAGAAATGGGAGGAAAGCGGACCTCTCGGAACTTCACTTTCACGT 246  
QY 181 GGTATTATGTTGATTCGATTCGTTGGCGCTCTGGCAATCTGTAGCTGTCTGTTTGTATC 240  
Db 247 GGTATTATGTTGATTCGATTCGTTGGCGCTCTGGCAATCTGTAGCTGTCTGTTTGTATC 306  
QY 241 TTGTTGACTATGAGGAAGTCTTAGGAAACTTAGGAATCTATGATGCTGATGATGGAG 300  
Db 307 TTGTTGACTATGAGGAAGTCTTAGGAAACTTAGGAATCTATGATGCTGATGATGGAG 366  
QY 301 ATTTTGTATGTTGATTCGTTGGCGCTCTGGCAATCTGTAGCTGTCTGTTTGTATC 360  
Db 367 ATTTTGTATGTTGATTCGTTGGCGCTCTGGCAATCTGTAGCTGTCTGTTTGTATC 426  
QY 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTG 420  
Db 427 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTG 486



QY 421 AGGAGAAACCCAGATATATCGAAGATGAAGCAAAAGCAAAATTCAGTCCCTTCCATG 480  
DB 487 AGGAGAAACCCAGATATATCGAAGATGAAGCAAAAGCAAAATTCAGTCCCTTCCATG 546  
QY 481 AAATGGGTACACGAGCAAAATGTTGAGGGAGAGAGCTTGCACAAAGAGATGACCCACAG 540  
DB 547 AAATGGGTACACGAGCAAAATGTTGAGGGAGAGAGCTTGCACAAAGAGATGACCCACAG 606  
QY 541 GAGAACCAACAAGAGAGATGATGAGTTCTTATGCGACTGATGTAGATGATAGATTG 600  
DB 607 GAGAACCAACAAGAGAGATGATGAGTTCTTATGCGACTGATGTAGATGATAGATTG 666  
QY 601 AGACCTGGAACCTGAAGTATCTATGAAGAACCGAGCATAGTTACCACTGGGAAGAGA 660  
DB 667 AGACCTGGAACCTGAAGTATCTATGAAGAACCGAGCATAGTTACCACTGGGAAGAGA 726  
QY 661 CAGTTTCAACAAGCTGTAATCAGGATATGGAAGAGATGATGTCTGAGCAGGAAATCCAG 720  
DB 727 CAGTTTCAACAAGCTGTAATCAGGATATGGAAGAGATGATGTCTGAGCAGGAAATCCAG 786  
QY 721 ATTCCAGTGAACCAAGTATGAGATCAAAAGATTGACCATCATATACATGATGATACAT 780  
DB 787 ATTCCAGTGAACCAAGTATGAGATCAAAAGATTGACCATCATATACATGATGATACAT 846  
QY 781 ACCAAGCTCTATGAGCAACAGCAGTATATGAACCTCTAGAAATGAGGGATGAAATCA 840  
DB 847 ACCAAGCTCTATGAGCAACAGCAGTATATGAACCTCTAGAAATGAGGGATGAAATCA 906  
QY 841 CAGAAGTAACTGCTCCCTGAGGATATCTGTGAAGATTCACAGGTAATTTGTAGAAG 900  
DB 907 CAGAAGTAACTGCTCCCTGAGGATATCTGTGAAGATTCACAGGTAATTTGTAGAAG 966  
QY 901 AAGTAAGCAATTTTCTGTGGAAGAACAGCAGGAGGATACCAACGAAACAAATAGAAA 960  
DB 967 AAGTAAGCAATTTTCTGTGGAAGAACAGCAGGAGGATACCAACGAAACAAATAGAAA 1026  
QY 961 CAGATGATCCAGAACAAAGCAAAAGTTAAAGAAAGAGCCTAAATTTAAATTAAT 1020  
DB 1027 CAGATGATCCAGAACAAAGCAAAAGTTAAAGAAAGAGCCTAAATTTAAATTAAT 1086  
QY 1021 TTGATAAGCAATTTAAAGCTGAATCTGTGCTGCAAGAAAATCTCGTAAAGGGGAAAA 1080  
DB 1087 TTGATAAGCAATTTAAAGCTGAATCTGTGCTGCAAGAAAATCTCGTAAAGGGGAAAA 1146  
QY 1081 TTGAGGAAGCACTGAATGCAATTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140  
DB 1147 TTGAGGAAGCACTGAATGCAATTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1206  
QY 1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGAGAGTAATGAGG 1200  
DB 1207 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGAGAGTAATGAGG 1266  
QY 1201 TGCTACGTGAGGACCATCGACCTACCAAGAGTGCCAGCCTACCTGATGTCCTGAG 1260  
DB 1267 TGCTACGTGAGGACCATCGACCTACCAAGAGTGCCAGCCTACCTGATGTCCTGAG 1326  
QY 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGCAACAAATTTCTAGTGCATATGA 1320  
DB 1327 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGCAACAAATTTCTAGTGCATATGA 1386  
QY 1321 GAGGTTCCCTGCTTACCTGACAGATAGTTAGTTCAACTATTTCCCAATGATCTCTCTAA 1380  
DB 1387 GAGGTTCCCTGCTTACCTGACAGATAGTTAGTTCAACTATTTCCCAATGATCTCTCTAA 1446  
QY 1381 AAAATGACCTTGGGTGGATACCTCTGTAGAGAGATGATGACATGCAAGAAAGTTT 1440  
DB 1447 AAAATGACCTTGGGTGGATACCTCTGTAGAGAGATGATGACATGCAAGAAAGTTT 1506  
QY 1441 ATGAAGAGTGTGAGTGA CACTTAATGATGGCTTGTGCTAAAGTCCATTTAGCTTCA 1500  
DB 1507 ATGAAGAGTGTGAGTGA CACTTAATGATGGCTTGTGCTAAAGTCCATTTAGCTTCA 1566  
QY 1501 TCCTGAAGGCAAGAACAAAATTTGCTGAGAGAGATCCCATATATTTAAAGAGGAAGATAGAT 1560

DB 1567 TCCTGAAGGCAAGAACAAAATTTGCTGAGAGATCCCATATTTAAAGAGGAATAGAT 1626  
QY 1561 CCGAGATCTCGGCATCTGATGATGGAGATTTTATTTCCACTGGGGATGCCATGAGA 1620  
DB 1627 CCGAGATCTCGGCATCTGATGATGGAGATTTTATTTCCACTGGGGATGCCATGAGA 1686  
QY 1621 GGGTTGGAAACAAGAGGCAATATAAGTGTGTATGAGCTTGGGCAAGAGAGGACACTTTG 1680  
DB 1687 GGGTTGGAAACAAGAGGCAATATAAGTGTGTATGAGCTTGGGCAAGAGAGGACACTTTG 1746  
QY 1681 CATCTGTCTGGCAACGCTCACTCTACAATGTGAATGGAATGAAAGCAGCAGCTTGGTGA 1740  
DB 1747 CATCTGTCTGGCAACGCTCACTCTACAATGTGAATGGAATGAAAGCAGCAGCTTGGTGA 1806  
QY 1741 CCCCCAAAAGAAACGGGCTTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAGTTAA 1800  
DB 1807 CCCCCAAAAGAAACGGGCTTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAGTTAA 1866  
QY 1801 TCCGAGATGAAGGCTTTCAGTGA TGGATAAAGCCAAAGGCTCTCTCTGCTGAGGATG 1860  
DB 1867 TCCGAGATGAAGGCTTTCAGTGA TGGATAAAGCCAAAGGCTCTCTCTGCTGAGGATG 1926  
QY 1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTTCACTGTGCGCAGCAAGGAAGAA 1920  
DB 1927 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTTCACTGTGCGCAGCAAGGAAGAA 1986  
QY 1921 ATGAAATGCTGCAAAAGGAGCTCTTAAACCTGTACCTTACTAGAAAGTTCCCCGAGA 1980  
DB 1987 ATGAAATGCTGCAAAAGGAGCTCTTAAACCTGTACCTTACTAGAAAGTTCCCCGAGA 2046  
QY 1981 CACAGAGATGCAAGAGAGAGACAGATCAAAATTTCCATGACACCCCGGACTCACGTGT 2040  
DB 2047 CACAGAGATGCAAGAGAGAGACAGATCAAAATTTCCATGACACCCCGGACTCACGTGT 2106  
QY 2041 GCGCGCACACAGGCCCCCAAACTGCGAGCTCCGAATGACACCTGGGCTTGGTATTTCCCA 2100  
DB 2107 GCGCGCACACAGGCCCCCAAACTGCGAGCTCCGAATGACACCTGGGCTTGGTATTTCCCA 2166  
QY 2101 AGGAAGCTGCAAGATTTCCATGTGCAACAGAGACAGAGCTGGGAGGAGGCAAGTGC 2160  
DB 2167 AGGAAGCTGCAAGATTTCCATGTGCAACAGAGACAGAGCTGGGAGGAGGCAAGTGC 2226  
QY 2161 TCATCTTTGATGACTCTTTTGACGAGGATGCGAGGATGCTCTATCTTTCCGCTGA 2220  
DB 2227 TCATCTTTGATGACTCTTTTGACGAGGATGCGAGGATGCTCTATCTTTCCGCTGA 2286  
QY 2221 TATTCATCTGAGTGTGTGGCATCCGGAATCTGACACCAACAGAGAGACGAGCCTTCCAG 2280  
DB 2287 TATTCATCTGAGTGTGTGGCATCCGGAATCTGACACCAACAGAGAGACGAGCCTTCCAG 2346  
QY 2281 CAATTTAGCATGAATTTATGCAAGCTTGGGAAACTCTGGAGAGA 2324  
DB 2347 CAATTTAGCATGAATTTATGCAAGCTTGGGAAACTCTGGAGAGA 2390

## RESULT 6

ADJ56257

ID ADJ56257 standard; cDNA; 5358 BP.

XX AC ADJ56257;

XX AC ADJ56257;

DT 06-MAY-2004 (first entry)

XX DE Bovine cDNA differentially expressed in MYCN activated cells seqID 63.

XX DE bovine; differential expression; transactivator; proto-oncogene;

KW neuroblastoma; small cell lung cancer; cytostatic; gene therapy; sa;

XX MYCN activated cell.

XX OS Bos taurus.

XX PN US2003119009-A1.



XX PD 26-JUN-2003.  
XX PF 25-FEB-2002; 2002US-00084817.  
XX PR 23-FEB-2001; 2001US-0270784P.  
XX PA (STUA//) STUART S G.  
XX PA (NUCH//) NUCHTERN J G.  
XX PA (PLON//) PLON S E.  
XX PA (SHOH//) SHOHET J M.  
XX Stuart SG, Nuchtern JG, Plon SE, Shohet JM;  
PI WPI; 2003-635698/60.  
XX New genes regulated by MYCN activation, useful in gene therapy,  
XX particularly for treating a subject with e.g. neuroblastoma or other  
XX cancers, or for diagnosing, staging or monitoring the treatment of the  
XX cancer.  
XX Claim 1; SEQ ID NO 63; 27pp; English.  
XX This invention relates to novel isolated cDNAs that are differentially  
XX expressed in MYCN activated cells. Specifically, it refers to  
XX polynucleotide sequences that exhibit differential expression patterns in  
XX cells activated by the transactivator MYCN, where MYCN is a proto-  
XX oncogene that is amplified in neuroblastoma cells and is common in small  
XX cell lung cancers. The present invention describes these cDNA molecules  
XX as useful for in hybridisation assays to detect expression of nucleic  
XX acids (or complementary nucleic acids) in a present in a given sample, as  
XX well as for screening assays by identifying molecules or compounds that  
XX specifically bind the cDNA as a ligand and modulate function or activity.  
XX Accordingly, these compositions exhibit cytostatic activity and can also  
XX be used for gene therapy purposes. This polynucleotide sequence is a cDNA  
XX that is differentially expressed in MYCN activated cells, given in an  
XX exemplification of the invention. NOTE: This sequence does not appear in  
XX the printed specification but has been obtained in electronic format from  
XX the US Patent Office at  
XX ftp://seqdata.uspto.gov/sequence.html?docID=20030119009.  
XX SQ Sequence 5358 BP; 1684 A; 1009 C; 1204 G; 1461 T; 0 U; 0 Other;  
Query Match 93.4%; Score 2171; DB 10; Length 5358;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2321; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 CGGACCGTGCATGGCCCGAGGTAGAAATGCCAAGAGCAGCGCGCAACAGCAGCAGCG 60  
DB 258 CGGACCGTGCATGGCCCGAGGTAGAAATGCCAAGAGCAGCGCGCAACAGCAGCAGCG 317  
QY 61 GCTCCGCGCAGGTAGCAGAGTGGCGGAGCAGCAGCGCGCGCGCGCGCGCGAGAGACAA 120  
DB 318 GCTCCGCGCAGGTAGCAGAGTGGCGGAGCAGCAGCGCGCGCGCGCGCGAGAGACAA 377  
QY 121 AGCATGGAGGACACAAGAAATGGGAGGAAAGGCGGACTCTCGGGAATTCATCTTCACGT 180  
DB 378 AGCATGGAGGACACAAGAAATGGGAGGAAAGGCGGACTCTCAGGAACTTCATCTTCACGT 437  
QY 181 GGTTATAGTGATGCAATGCTGGCGTCTGGACATCTGTAGCTGTCGTTTGGTTGATC 240  
DB 438 GGTTATAGTGATGCAATGCTGGCGTCTGGACATCTGTAGCTGTCGTTTGGTTGATC 497  
QY 241 TTGTTGACTATGAGGAAGTCTAGGAAACTAGGAATCTATGATGCTGATGTTGATGGAG 300  
DB 498 TTGTTGACTATGAGGAAGTCTAGGAAACTAGGAATCTATGATGCTGATGTTGATGGAG 557  
QY 301 ATTTTGAATGGATGATGCCAAAGTTTATTAGACCTTAAGAGAGATCTACTTCAGAGC 360  
DB 558 ATTTTGAATGGATGATGCCAAAGTTTATTAGAGCTTAAAGAGAGATCTACTTCAGAGC 617  
QY 361 CAGCAGTCCCGCCAGAGAGCGTGGAGCACAACACTGAGCCCGGAGGAGCGGTTCTCTGTGG 420

DB 618 CAGCAGTCCCGCCAGAGAGCGTGGAGCCACACACTGAGCCCGGAGGAGGTTCTGTGG 677  
QY 421 AGCAGAAACCCAGAAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG 480  
DB 678 AGCAGAAACCCAGAAATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATG 737  
QY 481 AAATGGTACACGAGAAACATGTTGAGGAGAGAACTTGCAACAAGAGATGCAACCCACAG 540  
DB 738 AAATGGTACACGAGAAACATGTTGAGGAGAGAACTTGCAACAAGAGATGCAACCCACAG 797  
QY 541 GAGAACCAACAAGAGAGATGATGATGTTCTTATGCGACTGATGATGATGATGATGATG 600  
DB 798 GAGAACCAACAAGAGAGATGATGATGTTCTTATGCGCACTGATGATGATGATGATG 857  
QY 601 AGACCTCTGGAACCTGAAGTATCTCATGAAGAAACCGAGCATGATGATGATGATGATG 660  
DB 858 AGACCTCTGGAACCTGAAGTATCTCATGAAGAAACCGAGCATGATGATGATGATGATG 917  
QY 661 CAGTTTCAAGAAGCTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAGAAATCCAG 720  
DB 918 CAGTTTCAAGAAGCTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAGAAATCCAG 977  
QY 721 ATTCCAGTGAACCAAGTAGTAGAAGATGAAAGATTGCACCATGATGATGATGATGATG 780  
DB 978 ATTCCAGTGAACCAAGTAGTAGAAGATGAAAGATTGCACCATGATGATGATGATGATG 1037  
QY 781 ACCAAGTCTATGAGGAACAAAGCAGTATATGAACCTCTAGAAAATGAAGGATGAATCA 840  
DB 1038 ACCAAGTCTATGAGGAACAAAGCAGTATATGAACCTCTAGAAAATGAAGGATGAATCA 1097  
QY 841 CAGAGTAACTGCTCCCTCGAGGATTAATCCTCTAGAAAGATTCCACAGGTAATTTGAGAAG 900  
DB 1098 CAGAGTAACTGCTCCCTCGAGGATTAATCCTCTAGAAAGATTCCACAGGTAATTTGAGAAG 1157  
QY 901 AAGTAAGCAATTTTCTCTGTGGAAGAAACAGCAGGAAGTACCACAGAAACAAATAGAAAA 960  
DB 1158 AAGTAAGCAATTTTCTCTGTGGAAGAAACAGCAGGAAGTACCACAGAAACAAATAGAAAA 1217  
QY 961 CAGATGATCCAGAAACAAAGCAAAAGTTAAGAAAAGAGAGCTTAACTTTTAAATAAT 1020  
DB 1218 CAGATGATCCAGAAACAAAGCAAAAGTTAAGAAAAGAGAGCTTAACTTTTAAATAAT 1277  
QY 1021 TTGATAAGACTATTAAAGCTGAACCTTGATGCTGAGAAAAAATCCGTAAGAGGAAAAA 1080  
DB 1278 TTGATAAGACTATTAAAGCTGAACCTTGATGCTGAGAAAAAATCCGTAAGAGGAAAAA 1337  
QY 1081 TTGAGGAAGCAAGTGAATGCAATTTAAAGAACTAGTACGCAATACCCCTCAGAGTCCAGAG 1140  
DB 1338 TTGAGGAAGCAAGTGAATGCAATTTAAAGAACTAGTACGCAATACCCCTCAGAGTCCAGAG 1397  
QY 1141 CAAGATATGGGAAGCGCGAGTGTGAGGATGATTTGGCTGAGAGAGAGAGAAATAGAGG 1200  
DB 1398 CAAGATATGGGAAGCGCGAGTGTGAGGATGATTTGGCTGAGAGAGAGAGAAATAGAGG 1457  
QY 1201 TGCTACGTGGAGCCATCGAGACCTTACCAAGAGGTGGCAGCCTACCTGATGCTCCCTGAG 1260  
DB 1458 TGCTACGTGGAGCCATCGAGACCTTACCAAGAGGTGGCAGCCTACCTGATGCTCCCTGAG 1517  
QY 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGCAGGCAACATTTCTAGTCTCATAGA 1320  
DB 1518 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGCAGGCAACATTTCTAGTCTCATAGA 1577  
QY 1321 GAGGTTCCCTGCTTACCTCGCAGAGATTTAGTTTCAACTATTTTCCCAATGATATCTTCTTAA 1380  
DB 1578 GAGGTTCCCTGCTTACCTCGCAGAGATTTAGTTTCAACTATTTTCCCAATGATATCTTCTTAA 1637  
QY 1381 AAAATGAACCTTGGCGTGGGATACCTCTCTGATAGGAGATTAATGACAAATGCAAGGAAGTTT 1440  
DB 1638 AAAATGAACCTTGGCGTGGGATACCTCTCTGATAGGAGATTAATGACAAATGCAAGGAAGTTT 1697  
QY 1441 ATGAGAGAGTGTGAGTGTGACCTAATGATGGCTTTGCTTAAAGTCCATTTATGCTTCA 1500  
DB 1698 ATGAGAGAGTGTGAGTGTGACCTAATGATGGCTTTGCTTAAAGTCCATTTATGCTTCA 1757

```
OY 1501 TCCTGAAGGCACAGACAAATTCCTGAGAGCATCCCATATTTAAAGAGGAGATAGAT 1560
DB 1758 TCCTGAAGGCACAGACAAATTCCTGAGAGCATCCCATATTTAAAGAGGAGATAGAT 1817
OY 1561 CCGAGATCCTGGCACTGATGATGGAGATTTATTTCCACTGGGGATGCCATGAGA 1620
DB 1818 CCGAGATCCTGGCACTGATGATGGAGATTTATTTCCACTGGGGATGCCATGAGA 1877
OY 1621 GGGTTGGGAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCACAGAGAGGACACTTTG 1680
DB 1878 GGGTTGGGAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCACAGAGAGGACACTTTG 1937
OY 1681 CATCTCTCTGGCAACGCTCACTCTACAAATGTAATGAGCTGAAGCAGAGCTTGGTGA 1740
DB 1938 CATCTCTCTGGCAACGCTCACTCTACAAATGTAATGAGCTGAAGCAGAGCTTGGTGA 1997
OY 1741 CCCCAAGAGAGAGGCTTACACAGAGTTAGTAACTTTAGAAAGAACTGGAAGTTAA 1800
DB 1998 CCCCAAGAGAGAGGCTTACACAGAGTTAGTAACTTTAGAAAGAACTGGAAGTTAA 2057
OY 1801 TCGAGATGAAGGCTTGCAGTATGATATAAGCCAAAGGCTCTTCTGCTGAGGATG 1860
DB 2058 TCGAGATGAAGGCTTGCAGTATGATATAAGCCAAAGGCTCTTCTGCTGAGGATG 2117
OY 1861 AAAAAGCTGAGGAAAGAGGCTTGGAGCCAGTTCAAGCTGTGGCAGCAAGGAAGAGAA 1920
DB 2118 AAAAAGCTGAGGAAAGAGGCTTGGAGCCAGTTCAAGCTGTGGCAGCAAGGAAGAGAA 2177
OY 1921 ATGAAATGCTTGCAGAGGAGCTTCTTAAAGCTGTACCTTACTAGAAAGTTCCCGAGA 1980
DB 2178 ATGAAATGCTTGCAGAGGAGCTTCTTAAAGCTGTACCTTACTAGAAAGTTCCCGAGA 2237
OY 1981 CAACAGATGAGAGGAGGACAGATCAATATTTCCATATGCAATGCAACCCGGGAGTCAAGTGT 2040
DB 2238 CAACAGATGAGAGGAGGACAGATCAATATTTCCATATGCAATGCAACCCGGGAGTCAAGTGT 2297
OY 2041 GGCAGACACAGGCCCCACAAATCGCAGGCTCCGAATGCACTGGGCTTGGTATCCCA 2100
DB 2298 GGCAGACACAGGCCCCACAAATCGCAGGCTCCGAATGCACTGGGCTTGGTATCCCA 2357
OY 2101 AGGAAGGCTGCAAGATTCGATGTGCAACAGAGCAGGACCTGGGAGGAAGCAAGGTGC 2160
DB 2358 AGGAAGGCTGCAAGATTCGATGTGCAACAGAGCAGGACCTGGGAGGAAGCAAGGTGC 2417
OY 2161 TCATCTTTGATGACTCTTTGAGCAGAGGTATGGAGATGCTCTATCTTTCCGCTGA 2220
DB 2418 TCATCTTTGATGACTCTTTGAGCAGAGGTATGGAGATGCTCTATCTTTCCGCTGA 2477
OY 2221 TATTCATGCTGGATGTGGCATCCGGAATGACACCAACAGAGAGCAGGCTTTCCAG 2280
DB 2478 TATTCATGCTGGATGTGGCATCCGGAATGACACCAACAGAGAGCAGGCTTTCCAG 2537
OY 2281 CAATTAGCATGAATTCATGCAAGCTTGGGAACTCTGAGAGA 2324
DB 2538 CAATTAGCATGAATTCATGCAAGCTTGGGAACTCTGAGAGA 2581
```

## RESULT 7

```
ABX04178
ID ABX04178 standard; cDNA; 2680 BP.
XX
AC ABX04178;
XX
DT 10-JAN-2003 (first entry)
XX
DE Human mRNA differentially expressed in mesenchymal cells #25.
XX
KW Human; ss; gene; skeletal growth; cartilage degeneration disorder;
KW chondroblastic phenotype; mesenchymal cell; cartilage formation;
KW bone formation; arthritis; osteoarthritis; rheumatoid arthritis;
KW gout arthritis; adjuvant arthritis; arthritis deformans; infectious
KW infectious arthritis; osteochondrosis; RDA; antiarthritis; osteopathic;
```

```
KW
XX
OS Homo sapiens.
XX
PN WO200271927-A2.
XX
PD 19-SBP-2002.
XX
PF 12-MAR-2002; 2002WO-US0077787.
XX
PR 12-MAR-2001; 2001US-0274980P.
XX
PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI Yates KE, Mizuno S, Glowacki J;
XX
WPI; 2002-723276/78.
XX
New nucleic acid molecules' capable of promoting chondrogenesis, useful
for diagnosing and treating cartilaginous tissue degeneration conditions,
e.g. osteoarthritis, rheumatoid arthritis, gout arthritis, or
osterochondrosis.
XX
Claim 33; Page 123; 153pp; English.
XX
The invention relates to new isolated nucleic acid molecule comprising a
nucleic acid molecule consisting of a gene differentially expressed in
cells undergoing differentiation from mesenchymal cell to a
chondroblastic phenotype, or hybridising under stringent conditions to
them (or their fragments). Also included are expression vectors,
transformed host cells, expressed polypeptides or peptide fragments
(which induce differentiation of a mesenchymal cell and may be used as an
immunogen), binding partners of the polypeptides, a method for
identifying an agent useful in modulating mesenchymal cell
differentiation induction activity of a molecule, a method of diagnosing
a condition characterized by aberrant expression of a nucleic acid
molecule or its expression product; a method for determining regression,
progression or onset of cartilaginous tissue degeneration condition in a
subject characterised by aberrant expression of a nucleic acid molecule
or its expression product, a method for treating a cartilaginous tissue
degeneration condition, a method for treating a subject to reduce the
risk of cartilaginous tissue degeneration condition developing in the
subject, a method for identifying a candidate agent for treating a
cartilaginous tissue degeneration condition, and a solid-phase nucleic
acid molecule array consisting essentially of a set of nucleic acid
molecule as cited above (or known from known genes shown to be
differentially expressed in developing mesenchymal cells using the
technique of representational difference analysis, RDA), its expression
products or fragments, fixed to a solid substrate. The nucleic acids,
polypeptides and agents are useful for treating cartilaginous tissue
degeneration conditions such as osteoarthritis, rheumatoid arthritis,
gout arthritis, adjuvant arthritis, arthritis deformans, infectious
arthritis or osteochondrosis. The present sequence is a cDNA from a
known gene differentially expressed in developing mesenchymal cells
XX
SQ Sequence 2680 BP; 854 A; 446 C; 575 G; 805 T; 0 U; 0 Other;
```

```
Query Match 38.5%; Score 895; DB 6; Length 2680;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 945; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 CGGACCGTGCATATGGCCAGCGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCG 60
DB 1 CGGACCGTGCATATGGCCAGCGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCG 60
OY 61 GCTCCGGCAGCGGTAGCAGAGTGGGGCAGCAGCAGCCCCGGGGCCCGAGAGAGACAA 120
DB 61 GCTCCGGCAGCGGTAGCAGAGTGGGGCAGCAGCAGCCCCGGGGCCCGAGAGAGACAA 120
OY 121 AGCATGGGAGCAGAGATGGGAGGAAGCGGACTCTCGGGAATTCATTCTTCAGGT 180
DB 121 AGCATGGGAGCAGAGATGGGAGGAAGCGGACTCTCAGGAATTCATTCTTCAGGT 180
```

QY 181 GGTATTGCTGATTCATGCTGGCGTCTGGACATCTGTAGCTGTGCTTTGGTTGATC 240  
DB |||||||  
QY 181 GGTATTGCTGATTCATGCTGGCGTCTGGACATCTGTAGCTGTGCTTTGGTTGATC 240  
DB |||||||  
QY 241 TTGTTGACTATGAGGAAGTTCTAGGAAATCTATGATGCTGTGATGGAG 300  
DB |||||||  
QY 241 TTGTTGACTATGAGGAAGTTCTAGGAAATCTATGATGCTGTGATGGAG 300  
DB |||||||  
QY 301 ATTTGATGCTGATGCTGCAAGTTTATTAGACTTAAAGAGAGATCTACTTCAGGC 360  
DB |||||||  
QY 301 ATTTGATGCTGATGCTGCAAGTTTATTAGACTTAAAGAGAGATCTACTTCAGGC 360  
DB |||||||  
QY 361 CAGAGTCCCGCCAGAGAGCTGAGCACACACTGAGCCCGAGGAGGTTCTCTGG 420  
DB |||||||  
QY 361 CAGAGTCCCGCCAGAGAGCTGAGCACACACTGAGCCCGAGGAGGTTCTCTGG 420  
DB |||||||  
QY 421 AGGCAGAACCCCGAGAAATTCGAAGATGAAGCAAAAGAAACAAATTCAGTCCATG 480  
DB |||||||  
QY 421 AGGCAGAACCCCGAGAAATTCGAAGATGAAGCAAAAGAAACAAATTCAGTCCATG 480  
DB |||||||  
QY 481 AAATGGTACGACAGAACTGTTGAGGAGAGACTTGCACAAAGAGATGACCCACAG 540  
DB |||||||  
QY 481 AAATGGTACGACAGAACTGTTGAGGAGAGACTTGCACAAAGAGATGACCCACAG 540  
DB |||||||  
QY 541 GAGAACCAACAGAGAGATGATGATGTTCTTATGGCGACTGATGATGATGATG 600  
DB |||||||  
QY 541 GAGAACCAACAGAGAGATGATGATGTTCTTATGGCGACTGATGATGATGATG 600  
DB |||||||  
QY 601 AGACCTCGAACCTGAAATCTCATGAAGAACCGAGCATAGTTACACGTGGAGAGA 660  
DB |||||||  
QY 601 AGACCTCGAACCTGAAATCTCATGAAGAACCGAGCATAGTTACACGTGGAGAGA 660  
DB |||||||  
QY 661 CAGTTTCACAGACTGTAATCAGGATATGGAAGATGATGCTGAGCAGGAAATCCAG 720  
DB |||||||  
QY 661 CAGTTTCACAGACTGTAATCAGGATATGGAAGATGATGCTGAGCAGGAAATCCAG 720  
DB |||||||  
QY 721 ATTCCAGTGAACCTGAGTAGAAGATGAAGATTCACCATGATACAGATGATGAAAT 780  
DB |||||||  
QY 721 ATTCCAGTGAACCTGAGTAGAAGATGAAGATTCACCATGATACAGATGATGAAAT 780  
DB |||||||  
QY 781 ACCAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA 840  
DB |||||||  
QY 781 ACCAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAATGAAGGATAGAAATCA 840  
DB |||||||  
QY 841 CAGAGTAACTGCTCCCTGAGGATAATCCTGTAGAGATTACAGAGTAATGTAGAAG 900  
DB |||||||  
QY 841 CAGAGTAACTGCTCCCTGAGGATAATCCTGTAGAGATTACAGAGTAATGTAGAAG 900  
DB |||||||  
QY 901 AAGTAAGCATTTTCTGTTGGAAGAACAGCAGGAAAGTACCACAGA 946  
DB |||||||  
QY 901 AAGTAAGCATTTTCTGTTGGAAGAACAGCAGGAAAGTACCACAGA 946  
DB |||||||  
RESULT 8  
ADP24208  
ID ADP24208 standard; cDNA; 2680 BP.  
XX  
XX ADP24208;  
AC  
XX  
XX 18-NOV-2004 (first entry)  
DT  
DE PRO polypeptide encoding cDNA SEQ ID NO:1386.  
XX  
XX ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;  
KW immunosuppressive; osteopathic; antidiabetic; dermatological;  
KW antipsoriatic; antiallergic; antiasthmatic; hepatotropic; respiratory;  
KW gene therapy; immune system.  
XX  
XX Unidentified.  
OS  
XX WO2004041170-A2.  
PN  
XX  
XX 21-MAY-2004.  
PD

XX 30-OCT-2003; 2003WO-US034312.  
PF  
XX 01-NOV-2002; 2002US-0423394P.  
PR  
XX (GETH ) GENENTECH INC.  
PA  
XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;  
PI Wu TD;  
PI WPI; 2004-419628/39.  
DR P-PSDB; ADP42409.  
DR  
XX New PRO polypeptides and polynucleotides, useful for treating e.g.  
PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated  
PT renal disease, or demyelinating diseases of the central or peripheral  
PT nervous system.  
XX  
XX Claim 1; SEQ ID NO 1386; 2940pp; English.  
XX  
XX The invention relates to a novel isolated nucleic acid and the PRO  
CC polypeptide encoded by it. A protein of the invention has  
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,  
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,  
CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide  
CC of the invention may have a use in gene therapy. The PRO polypeptide, its  
CC agonist, antagonist, or antibody that specifically binds to the  
CC polypeptide is useful for treating an immune related disorder such as  
CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,  
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an  
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic  
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune  
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal  
CC disease, a demyelinating disease of the central or peripheral nervous  
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,  
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary  
CC disease, infectious or autoimmune chronic active hepatitis, primary  
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,  
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's  
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin  
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic  
CC disease, asthma, allergic rhinitis, atopic dermatitis, food  
CC hypersensitivity, urticaria, an immunologic disease of the lung,  
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity  
CC pneumonitis, a transplantation associated disease, graft rejection or  
CC graft-versus-host disease. The present sequence encodes a PRO protein of  
XX the invention.  
SQ Sequence 2680 BP; 854 A; 446 C; 575 G; 805 T; 0 U; 0 Other;  
Query Match 38.5%; Score 895; DB 13; Length 2680;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 945; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 CGGACCGTGCAATGGCCCGAGCGTAAGATGCCAAGAGCAGCGCAACAGCAGCAGCAGCG 60  
DB |||||||  
QY 1 CGGACCGTGCAATGGCCCGAGCGTAAGATGCCAAGAGCAGCGCGCAACAGCAGCAGCAGCG 60  
DB |||||||  
QY 61 GCTCCGGCAGCGGTAGCACGAGTGGGCGACAGCAGCGCCCGGGGCCCGGAGAGACAA 120  
DB |||||||  
QY 61 GCTCCGGCAGCGGTAGCACGAGTGGGCGACAGCAGCGCCCGGGGCCCGGAGAGACAA 120  
DB |||||||  
QY 121 AGCATGGAGGACACAGAAATGGGAGAAAGCGGAACTCTCGGGAACCTTCATCTTCAGCT 180  
DB |||||||  
QY 121 AGCATGGAGGACACAGAAATGGGAGAAAGCGGAACTCTCGGGAACCTTCATCTTCAGCT 180  
DB |||||||  
QY 181 GGTATTGCTGATTCATGCTGGCGTCTGGACATCTGTAGCTGTGCTTTGGTTGATC 240  
DB |||||||  
QY 181 GGTATTGCTGATTCATGCTGGCGTCTGGACATCTGTAGCTGTGCTTTGGTTGATC 240  
DB |||||||  
QY 241 TTGTTGACTATGAGGAAGTTCTAGGAAATCTATGATGCTGATGCTGATGAGGAG 300  
DB |||||||  
QY 241 TTGTTGACTATGAGGAAGTTCTAGGAAATCTATGATGCTGATGCTGATGAGGAG 300  
DB |||||||

QY 301 ATTTGATGTGGATGATGCCAAAGTTTTTATTTAGGACTTAAAGAGAGATCTACTTCAGAGC 360  
Db |||||  
QY 301 ATTTGATGTGGATGATGCCAAAGTTTTTATTTAGGACTTAAAGAGAGATCTACTTCAGAGC 360  
Db |||||  
QY 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGG 420  
Db |||||  
QY 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGG 420  
Db |||||  
QY 421 AGCAGACCCAGAGATATCGAAGATGAACAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
Db |||||  
QY 421 AGCAGACCCAGAGATATCGAAGATGAACAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
Db |||||  
QY 481 AATGTGACACGAGAAATCTTTGAGGAGAGAGCTTGCAACAAGAGATGAGCCACAG 540  
Db |||||  
QY 481 AATGTGACACGAGAAATCTTTGAGGAGAGAGCTTGCAACAAGAGATGAGCCACAG 540  
Db |||||  
QY 541 GAGAACCAACAGAGAGATGATGAGTTTCTTATGCGGACTGATGTAGATGATGATTTG 600  
Db |||||  
QY 541 GAGAACCAACAGAGAGATGATGAGTTTCTTATGCGGACTGATGTAGATGATGATTTG 600  
Db |||||  
QY 601 AGACCTGGAACCTGAAGTATCTCATGAGAAACCGAGCATAGTTTACACGTGGAAGA 660  
Db |||||  
QY 601 AGACCTGGAACCTGAAGTATCTCATGAGAAACCGAGCATAGTTTACACGTGGAAGA 660  
Db |||||  
QY 661 CAGTTTCAACAGCTGATATCAGGATATGAGAGATGATGTTGAGCAGGAAATCCAG 720  
Db |||||  
QY 661 CAGTTTCAACAGCTGATATCAGGATATGAGAGATGATGTTGAGCAGGAAATCCAG 720  
Db |||||  
QY 721 ATTCCAGTGAACCAAGTATGAGAGATGAGTTGACCATGATACAGATGATTAACAT 780  
Db |||||  
QY 721 ATTCCAGTGAACCAAGTATGAGAGATGAGTTGACCATGATACAGATGATTAACAT 780  
Db |||||  
QY 781 ACCAAGTCTATGAGAAACAGCAGTATATGAACCTTAGAAATGAAGGATGAATCA 840  
Db |||||  
QY 781 ACCAAGTCTATGAGAAACAGCAGTATATGAACCTTAGAAATGAAGGATGAATCA 840  
Db |||||  
QY 841 CAGAGTAACTGCTCCCGCTGAGGATATCTGTAAGATTCAGGTTAATGTTGAAG 900  
Db |||||  
QY 841 CAGAGTAACTGCTCCCGCTGAGGATATCTGTAAGATTCAGGTTAATGTTGAAG 900  
Db |||||  
QY 901 AAGTAAGATTTTCTGTGGAAGAACAGACAGAGAGTACCACAGA 946  
Db |||||  
QY 901 AAGTAAGATTTTCTGTGGAAGAACAGACAGAGAGTACCACAGA 946  
Db |||||

## RESULT 9

AAZ23609  
ID AAZ23609 standard; cDNA; 2442 BP.

AC AAZ23609;

DT 06-JAN-2000 (first entry)

DE Human labyrinthin cDNA.

XW Labyrinthin; human; cancer; marker; antigen; detection; antibody;  
MCA 44-3A6; diagnostic; vaccine; treatment; adenocarcinoma; ds.

XX Homo sapiens.

OS Key Location/Qualifiers  
PH 70..837  
FT /\*tag= a  
FT /product= "lab"

XX WO947683-A1.

XX 23-SEP-1999.

XX 11-MAR-1999; 99WO-US0005365.

XX 17-MAR-1998; 98US-00040485.

XX  
PA

(RADO/) RADOSEVICH J A.

XX Radosevich JA;

XX WPI; 1999-580307/49.

XX P-PSDB; AAY33642.

XX

Novel Labyrinthin polynucleotides and polypeptides used as a diagnostic marker for cancer and in anticancer vaccines.

Claim 1; Fig 1; 34pp; English.

This invention describes a novel polynucleotide and polypeptide sequence for the antigen detected by the antibody MCA 44-3A6. This antigen is designated Labyrinthin (Lab). Antibodies directed to the Labyrinthin (Lab) protein are useful for the detection of Lab. The Lab protein is useful in diagnostic assays for cancer, e.g. to monitor the presence and amount of antibodies (this method is especially useful for cancer cells that have the Lab marker). As the Lab gene is not tissue-specific, it will detect cancer regardless of which organ it occurs in. Peptides derived from Lab are used in the preparation of vaccines to prevent human cancers and/or to treat humans with cancer. Antibody MCA 44-3A6 is able to differentiate antigens associated with adenocarcinomas. However, the sequence of the antigen detected by this antibody has not been elucidated in the prior art. Determination of the polypeptide and polynucleotide sequence of this antigen would enhance its usefulness in cancer diagnosis, treatment and prevention. The present invention discloses the sequence of the antigen recognized by the MCA 44-3A6 antibody. This sequence encodes the human Lab protein described in the method of the invention

SQ Sequence 2442 BP; 781 A; 391 C; 512 G; 758 T; 0 U; 0 Other;

Query Match 27.5%; Score 640; DB 2; Length 2442;

Best Local Similarity 99.9%; Pred. No. 6.5e-287;

Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 162 GGAATCTTCATCTTCACTGGTTTATGGTGGTTCATTTGGGCGTCTGGACATCTGTA 221

Db 46 GGAATCTTCATCTTCACTGGTTTATGGTGGTTCATTTGGGCGTCTGGACATCTGTA 105

QY 222 GCTGCTGTTGTTTGTATCTTGTGACTATGAGAGTCTTCTGGAAGATGAGTATCTAT 281

Db 106 GCTGCTGTTGTTTGTATCTTGTGACTATGAGAGTCTTCTGGAAGATGAGTATCTAT 165

QY 282 GATCTGATGTTGATGAGATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 341

Db 166 GATCTGATGTTGATGAGATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 225

QY 342 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCC 401

Db 226 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCC 285

QY 402 GAGGAGCAGGTTCTCTGTGAGGAGAGCCAGAGATATCGAAGATCAAGCAAGAACAA 461

Db 286 GAGGAGCAGGTTCTCTGTGAGGAGAGCCAGAGATATCGAAGATCAAGCAAGAACAA 345

QY 462 ATTCAAGTCCCTTCTCCATGAATGTTGATGATGATGATGATGATGATGATGATGATGAT 521

Db 346 ATTCAAGTCCCTTCTCCATGAATGTTGATGATGATGATGATGATGATGATGATGATGAT 405

QY 522 CAGAGATGAGGAGCCAGAGAGAACCAACAGAGAGATGATGATGATGATGATGATGATGAT 581

Db 406 CAGAGATGAGGAGCCAGAGAGAACCAACAGAGAGATGATGATGATGATGATGATGATGAT 465

QY 582 GATGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 641

Db 456 GATGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 525

QY 642 AGTTACACGTGGAAGAGACAGTTCACAGAGCTGTAATCAGAGATGATGATGATGATGATG 701

Db 526 AGTTACACGTGGAAGAGACAGTTCACAGAGCTGTAATCAGAGATGATGATGATGATGATG 585



PR 13-DEC-2000; 2000US-00736457.  
XX (BANG// BANGUR C S.  
PA (FANG// FANGER G R.  
PA (WANG// WANG A.  
PA (WANG// WANG T.  
PA (SWIT// SWITZER A P.  
PA (MCNE// MCNEILL P D.  
PA (CLAP// CLAPPER J D.  
XX  
PI Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, McNeill PD;  
PI Clapper JD;  
XX  
XX WPI; 2003-352750/33.  
DR P-PSDB; ABU69516.  
XX  
PT Novel lung cancer polynucleotide encoding lung cancer protein, useful for  
PT detecting the presence of lung cancer in a patient, and in pharmaceutical  
PT compositions, e.g. vaccines, for treating lung cancer.  
XX  
XX Example 5; Page: 72pp; English.

XX The invention relates to a polynucleotide encoding a lung tumour protein,  
CC comprising a sequence selected from any of the 14 sequences mentioned in  
CC the specification, or a sequence (S2) mentioned in specification,  
CC complement of S1, sequences consisting of at least 20 contiguous residues  
CC of S1, sequences that hybridise to S1, sequences having 75% preferably  
CC 90%, identity to S1, or degenerate variants of S1. Also included are an  
CC isolated polypeptide (comprising a sequence (S3) selected from any one of  
CC the 4 amino acid sequences mentioned in the specification, a sequence  
CC encoded by the polynucleotide, or sequences having at least 70%,  
CC preferably 90%, identity to a sequence encoded by the polynucleotide), an  
CC expression vector comprising the polynucleotide operably linked to an  
CC expression control sequence, a host cell transformed or transfected with  
CC the vector, an isolated antibody (or its antigen-binding fragment) that  
CC specifically binds to the polypeptide, detecting the presence of a cancer  
CC in a patient, a fusion protein comprising the polypeptide, an  
CC oligonucleotide that hybridises to S1 under moderately stringent  
CC conditions, stimulating and/or expanding T cells with the polynucleotide, a  
CC protein (comprising contacting T cells with the polynucleotide, protein  
CC or antigen-presenting cells, under conditions and for a time sufficient  
CC to permit the stimulation and/or expansion of T cells) and inhibiting the  
CC development of a cancer in a patient (by incubating CD4<sup>+</sup> and/or CD8<sup>+</sup> T  
CC cells isolated from a patient with the polynucleotide, protein or antigen  
CC presenting cells that express the polynucleotide, such that T cells  
CC proliferate, administering to the patient an effective amount of the  
CC the patient, T cells, and thus inhibiting the development of a cancer in  
CC the patient. The polynucleotide, protein and cells are useful in a  
CC composition for stimulating an immune response in a patient, and for  
CC treating a cancer in a patient (particularly lung cancer). The  
CC oligonucleotide is useful for determining the presence of a cancer in a  
CC patient. The protein and oligonucleotides are useful in pharmaceutical  
CC compositions, e.g. vaccines. The polynucleotide is also useful as a probe  
CC or primer for nucleic acid hybridisation, and in the design and  
CC preparation of ribozyme molecules for inhibiting expression of tumour  
CC polypeptides and proteins in tumour cells. An amplified portion of the  
CC polynucleotide is useful for isolating a full-length gene from a suitable  
CC library. The present sequence is a cDNA (full length, extended or  
CC partial) isolated from a library derived from lung tumour/cancer cells.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from the  
CC USPTO at seqdata.uspto.gov/sequence.html?DocId=20020197669

XX  
SQ Sequence 2442 BP; 781 A; 391 C; 512 G; 758 T; 0 U; 0 Other;  
Query Match 27.58; Score 640; DB 8; Length 2442;  
Best Local Similarity 99.94; Pred. No. 6.5e-287;  
Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 162 GGAATTCATCTTCACGCTGTTATGGTGATGTCATGCTGGCGCTGGACATCTGTA 221  
DB  
46 GGAATTCATCTTCACGCTGTTATGGTGATGTCATGCTGGCGCTGGACATCTGTA 105

QY 222 GCTGCTGTTGGTTGATCTTGTTGACTATGAGGAGTCTTAGGAAAACCTAGGAATCTAT 281  
DB 106 GCTGCTGTTGGTTGATCTTGTTGACTATGAGGAGTCTTAGGAAAACCTAGGAATCTAT 165  
QY 282 GATGCTGATGGTGATGAGGATTTTGTATGTGGATGATGCCAAAGTTTATTAGGACTTAAA 341  
DB 166 GATGCTGATGGTGATGAGGATTTTGTATGTGGATGATGCCAAAGTTTATTAGGACTTAAA 225  
QY 342 GAGAGATCTACTTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCC 401  
DB 226 GAGAGATCTACTTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCC 285  
QY 402 GAGGAGCAGGTTCTCTGTGGAGGAGAGCCAGATATCGAAGATGAGCAAGAGACAA 461  
DB 286 GAGGAGCAGGTTCTCTGTGGAGGAGAGCCAGATATCGAAGATGAGCAAGAGACAA 345  
QY 462 ATTCACTCCCTTCTCGATGAATGCTACAGCGAGAGACATGTTGAGGAGAGAGACTTGCAG 521  
DB 346 ATTCACTCCCTTCTCGATGAATGCTACAGCGAGAGACATGTTGAGGAGAGAGACTTGCAG 405  
QY 522 CAAGAAGATGAGCCACAGAGAGAACCAACCAAGAGAGATGATGATTTCTTATGGCGACT 581  
DB 406 CAAGAAGATGAGCCACAGAGAGAACCAACCAAGAGAGATGATGATTTCTTATGGCGACT 465  
QY 582 GATGTAGATGATGATTTGAGACCCCTGGAACTCTGAAGTATCTCATGAAGAAACCGAGCAT 641  
DB 466 GATGTAGATGATGATTTGAGACCCCTGGAACTCTGAAGTATCTCATGAAGAAACCGAGCAT 525  
QY 642 AGTTACCACTGGAGAGAGACAGTTTTCACAGAGCTCTTAATCAGGATATGAGGAGAGATGATG 701  
DB 526 AGTTACCACTGGAGAGAGACAGTTTTCACAGAGCTCTTAATCAGGATATGAGGAGAGATGATG 585  
QY 702 TCTGAGCAGGAAATCCAGATTCAGATGAAACAGATGATGAGATGAGAGATGATG 761  
DB 586 TCTGAGCAGGAAATCCAGATTCAGATGAAACAGATGATGAGATGAGAGATGATG 645  
QY 762 GATACAGATGATGATTAACATACCAAGTCTATGAGGAGAACAGCAGTATATGAACCTCTAGAA 821  
DB 646 GATACAGATGATGATTAACATACCAAGTCTATGAGGAGAACAGCAGTATATGAACCTCTAGAA 705  
QY 822 AATGAAGGATAGAAATTCACAGAGTAACTG 852  
DB 706 AATGAAGGATAGAAATTCACAGAGTAACTG 736.

## RESULT 12

ACR03258  
ID ACR03258 standard; cDNA; 2442 BP.

XX AC ACR03258;

XX 22-MAY-2003 (first entry)

DE Lung cancer therapy and diagnosis associated cDNA #1745.

XX Lung cancer; cytostatic; vaccine; gene therapy; cancer; gene; ss.

XX Homo sapiens.

OS

PN US2002172952-A1.

XX 21-NOV-2002.

XX 10-JUL-2001; 2001US-00902941.

XX 30-JUN-1999; 99US-00346492.

PR 15-OCT-1999; 99US-00419356.

PR 17-DEC-1999; 99US-00468867.

PR 30-DEC-1999; 99US-00476300.

PR 06-MAR-2000; 2000US-00519642.

PR 22-MAR-2000; 2000US-00533077.

PR 10-APR-2000; 2000US-00546259.

PR 27-APR-2000; 2000US-00560406.



PR 05-JUN-2000; 2000US-00589184.  
PR 11-JUL-2000; 2000US-00614124.  
PR 29-AUG-2000; 2000US-00651563.  
PR 08-SEP-2000; 2000US-00658824.  
PR 26-SEP-2000; 2000US-00671325.  
PR 06-OCT-2000; 2000US-00677419.  
PR 30-OCT-2000; 2000US-00702705.  
PR 13-DEC-2000; 2000US-00736457.  
PR 03-MAY-2001; 2001US-00849626.  
XX (CORI-) CORIXA CORP.  
XX  
XX Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW;  
PI Durham M, Carter D, Fanger GR, Vedvick TS, Bangur CS, McNabb A;  
XX  
XX WPI; 2003-328427/31.  
XX  
XX New polynucleotide, useful for preparing a composition for treating or  
PT inhibiting development of cancer, e.g. lung cancer.  
XX  
XX Example 5; SEQID NO 1791; 82pp; English.  
XX  
XX The invention describes an isolated polynucleotide comprising one of 32  
CC sequences, complement or degenerate variants of them. The polynucleotide  
CC is useful for preparing a composition e.g. a vaccine or for gene therapy,  
CC for treating or inhibiting development of cancer, e.g. lung cancer. This  
CC sequence represents a polynucleotide associated with the compositions and  
CC methods for the therapy and diagnosis of lung cancer  
XX  
XX Sequence 2442 BP; 781 A; 391 C; 512 G; 758 T; 0 U; 0 Other;  
SQ

Query Match 27.5%; Score 640; DB 8; Length 2442;  
Best Local Similarity 99.9%; Pred. No. 6.5e-287;  
Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 162 GGAACCTTCATCTTCACGTGGTTTATGGTGAATTCGATGCTGGCGCTGGACATCTGTA 221  
DB 46 GGAACCTTCATCTTCACGTGGTTTATGGTGAATTCGATGCTGGCGCTGGACATCTGTA 105  
QY 222 GCTGCTGTTGTTGATCTTGTACATATGAGGAAGTCTTAGGAAACTAGGAATCTAT 281  
DB 106 GCTGCTGTTGTTGATCTTGTACATATGAGGAAGTCTTAGGAAACTAGGAATCTAT 165  
QY 282 GATGCTGATGGTGATGGAGATTTTCATGTGGATGATGCGCAAGTTTATTAGGACTTAAA 341  
DB 166 GATGCTGATGGTGATGGAGATTTTCATGTGGATGATGCGCAAGTTTATTAGGACTTAAA 225  
QY 342 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCC 401  
DB 226 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCC 285  
QY 402 GAGGACAGGTTCTCTGAGGACGAGAACCCCGAGATATCGAAGATGAAGCAAAAGACAA 461  
DB 286 GAGGACAGGTTCTCTGAGGACGAGAACCCCGAGATATCGAAGATGAAGCAAAAGACAA 345  
QY 462 ATTCACTCCCTTCCATCAAAATGGTACACGAGAACATGTTGAGGAGAACACTTGCAA 521  
DB 346 ATTCACTCCCTTCCATCAAAATGGTACACGAGAACATGTTGAGGAGAACACTTGCAA 405  
QY 522 CAAGAAGATGGACCCACAGAGAACCAACAAAGAGGATGATGAGTTTCTTATGGGACT 581  
DB 406 CAAGAAGATGGACCCACAGAGAACCAACAAAGAGGATGATGAGTTTCTTATGGGACT 465  
QY 582 GATGTAGATGATAGATTTTCAGACCCCTGGAACTTCAATGAGTATCTCATGAAGAAACCGAGCAT 641  
DB 466 GATGTAGATGATAGATTTTCAGACCCCTGGAACTTCAATGAGTATCTCATGAAGAAACCGAGCAT 525  
QY 642 AGTTACACCGTGAAGAGACAGTTTTCACAGACTGTAATCAGGATATGAAGAGATGATG 701  
DB 526 AGTTACACCGTGAAGAGACAGTTTTCACAGACTGTAATCAGGATATGAAGAGATGATG 585  
QY 702 TCTGAGCAGGAAATTCAGATTCAGTGAACCACTAGTAGAGATCAAGAGATTGACCAT 761

DB 586 TCTGAGCAGGAAAATCCAGATTCCAGTGAACACAGTAGTAGAAGATGAAGATTGCACCAT 645  
QY 762 GATACAGATGATGTAAACATACCAAGTCTATGAGGAACAGCAGATATATGAACCTCTAGAA 821  
DB 646 GATACAGATGATGTAAACATACCAAGTCTATGAGGAACAGCAGATATATGAACCTCTAGAA 705  
QY 822 AATGAGGAGGATGAATAATCACAGAAGTAACTG 852  
DB 706 AATGAGGAGGATGAATAATCACAGAAGTAACTG 736  
RESULT 13  
ADH47310  
ID ADH47310 standard; cDNA; 2442 BP.  
XX  
XX ADH47310;  
AC  
XX 25-MAR-2004 (first entry)  
DT Human lung tumour cDNA clone, SEQ ID NO 1791.  
XX  
DE Human lung tumour cDNA clone, SEQ ID NO 1791.  
XX  
XX lung cancer; tumour; cytostatic; immunostimulant; vaccine; gene therapy;  
KW human; clone; ss.  
KW  
XX Homo sapiens.  
OS  
XX WO2003037267-A2.  
PN  
XX 08-MAY-2003.  
DD  
XX 28-OCT-2002; 2002WO-US034777.  
PF  
XX 29-OCT-2001; 2001US-00017754.  
PR  
XX 28-MAR-2002; 2002US-00113872.  
PR  
XX (CORI-) CORIXA CORP.  
PA  
XX Henderson RA, Wang T, Watanabe Y, Kalos MD, Sleath PR; Vedvick TS;  
PI Johnson JC, Retter MW, Durham M, Carter D, Fanger GR, Vedvick TS;  
PI Bangur CS, McNabb A;  
PI  
XX WPI; 2003-468346/44.  
DR  
XX New polypeptides and encoding polynucleotides, useful for diagnosing,  
PT preventing and/or treating lung cancer.  
PT  
XX Example 5; SEQ ID NO 1791; 258pp; English.  
PS  
XX The invention relates to novel compositions and methods for the therapy  
CC and diagnosis of cancer, particularly lung cancer. The compositions  
CC comprise one or more lung tumour polypeptides, immunogenic portions  
CC thereof, polynucleotides that encode such polypeptides, antigen presenting  
CC cells that express such polypeptides, and T cells that are specific for  
CC cells expressing such polypeptides. The novel compositions have  
CC cytostatic and immunostimulant activity. The lung tumour antigens can be  
CC used in the creation of a vaccine. The polynucleotides that encode the  
CC lung tumour polypeptides can be used in gene therapy to help in the  
CC treatment of lung tumours. This polynucleotide sequence represents a  
CC human lung tumour cDNA clone of the invention. This sequence was not  
CC shown in the specification. It has been taken from a World Intellectual  
CC Property Organization CD ROM supplied with the specification.  
XX  
XX Sequence 2442 BP; 781 A; 391 C; 512 G; 758 T; 0 U; 0 Other;  
SQ

Query Match 27.5%; Score 640; DB 10; Length 2442;  
Best Local Similarity 99.9%; Pred. No. 6.5e-287;  
Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 162 GGAACCTTCATCTTCACGTGGTTTATGGTGAATTCGATGCTGGCGCTGGACATCTGTA 221  
DB 46 GGAACCTTCATCTTCACGTGGTTTATGGTGAATTCGATGCTGGCGCTGGACATCTGTA 105  
QY 222 GCTGCTGTTGTTGATCTTGTACATATGAGGAAGTCTTAGGAAACTAGGAATCTAT 281







XX 12-OCT-2001; 2001US-0329689P.  
PR 25-OCT-2001; 2001US-0335703P.  
PR 09-NOV-2001; 2001US-0348887P.  
PR 28-NOV-2001; 2001US-0334145P.  
PR 06-DEC-2001; 2001US-0337451P.  
PR 14-DEC-2001; 2001US-0340584P.  
XX (INCY-) INCYTE GENOMICS INC.  
XX Rankumar J., Gorvad AE, Baughn MR, Emerling BM, Yang J, Lee SY;  
PI Tran UK, Becha SD, Duggan BM, Lee EA, Griffin JA, Li JX;  
PI Sprague WW, Hafalia AJA, Chawla NK, Lehr-Mason PM, Kable AE, Yue H;  
PI Marquis JP, Yao MG, Richardson TW, Tang TV, Jin P, Chien D;  
PI Bhatia U, Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;  
XX MPI; 2003-430274/40.  
DR P-PSDB; ABU92053.  
XX New human protein modification and maintenance molecules (PMMM), useful  
PI for diagnosing, treating and preventing diseases or conditions associated  
PT with the aberrant PMMM expression e.g. cancer, atherosclerosis, or  
PT infections.  
XX Claim 5; Page 306; 311pp; English.  
XX The present invention relates to the isolation of human protein  
CC modification and maintenance molecules (PMMM), and the polynucleotide  
CC sequences encoding them. A total of 40 PMMM polypeptides (designated PMMM  
CC -1 to PMMM-40) are disclosed. The sequences of the invention are useful  
CC for diagnosing a condition or disease associated with the expression of  
CC PMMM in a subject, preparing a polyclonal or monoclonal antibody, and  
CC generating an expression profile of a sample containing the  
CC polynucleotides. The diseases or conditions associated with decreased  
CC expression or overexpression of PMMM are cell proliferation disorders  
CC (e.g. cancer, atherosclerosis), neurological disorders (e.g. epilepsy,  
CC Huntington's disease, stroke), immune/inflammatory disorders, (e.g. AIDS,  
CC allergies), developmental disorders (e.g. hypothyroidism, Cushing's  
CC syndrome), gastrointestinal or epithelial disorders, and infections. The  
CC PMMM polypeptides or their fragments are useful in screening compounds  
CC for effectiveness as agonists or antagonists of the polypeptides, or in  
CC altering the expression of the target polynucleotide and compounds that  
CC specifically bind to, or modulate the activity of the polypeptide.  
CC ACA92416-ACA92455 encode the human PMMM polypeptides of the invention  
XX  
SQ Sequence 2648 BP; 844 A; 420 C; 545 G; 839 T; 0 U; 0 Other;  
Query Match 21.6%; Score 502; DB 10; Length 2648;  
Best Local Similarity 100.0%; Pred. No. 1e-222;  
Matches 502; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 162 GGAACCTTCATTCTTCCAGTGGTTTATGGTGGTATGCGTGGCGCTCGACATCTGTA 221  
DB 179 GGAACCTTCATTCTTCCAGTGGTTTATGGTGGTATGCGTGGCGCTCGACATCTGTA 238  
QY 222 GCTGTCGTTTGGTTGATCTTGTGACTATGAGGAAGTTCTAGCAAACTAGCAATCTAT 281  
DB 239 GCTGTCGTTTGGTTGATCTTGTGACTATGAGGAAGTTCTAGCAAACTAGCAATCTAT 298  
QY 282 GATCTGATGCTGATGAGATTTTGTGATGCTGATGATGCGCAAGTTTATTAGCACTTAA 341  
DB 299 GATGCTGATGCTGATGAGATTTTGTGATGCTGATGATGCGCAAGTTTATTAGCACTTAA 358  
QY 342 GAGAGATCTACTTCAGAGCAGCAGTCCCGCCAGAGAGGCTGAGCCACACTGAGCCC 401  
DB 359 GAGAGATCTACTTCAGAGCAGCAGTCCCGCCAGAGAGGCTGAGCCACACTGAGCCC 418  
QY 402 GAGGAGCAGGTTCTCTGAGGAGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAACAA 461  
DB 419 GAGGAGCAGGTTCTCTGAGGAGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAACAA 478  
QY 462 ATTTCAGTCCCTTCTCCATGAATGGTTACACGAGCAACATGTTGAGGAGAGACTTGCAA 521

Db 479 ATTTCAGTCCCTTCTCCATGAATGGTTACACGAGCAACATGTTGAGGAGGAGAGACTTGCAG 538  
QY 522 CAAGAAGATGAGCCCAAGAGAGAACCAACCAAGAGGATGATGATGTTCTTATGGCGACT 581  
Db 539 CAAGAAGATGAGCCCAAGAGAGAACCAACCAAGAGGATGATGATGTTCTTATGGCGACT 598  
QY 582 GATGTAGATGATGATTTGAGACCCCTGGAACTCTGAACTGATCTCTATGAGAAACCGAGCAT 641  
Db 599 GATGTAGATGATGATTTGAGACCCCTGGAACTCTGAACTGATCTCTATGAGAAACCGAGCAT 658  
QY 642 AGTTACACGCTGGAGAGACAG 663  
Db 659 AGTTACACGCTGGAGAGACAG 680  
RESULT 17  
ACH47067  
ID ACH47067 standard; cDNA; 502 BP.  
XX  
AC ACH47067;  
XX AC ACH47067;  
DT 13-OCT-2003 (first entry)  
XX  
XX Human infant brain cDNA #1130.  
XX  
XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
KW genome mapping; biodiversity; genetic disorder.  
XX Homo sapiens.  
OS  
XX US2003073623-A1.  
PN  
PD 17-APR-2003.  
XX  
XX 30-JUL-2001; 2001US-00918995.  
PF  
XX 30-JUL-2001; 2001US-00918995.  
PR  
XX (DRMA/) DRMANAC R T.  
PA (LABA/) LABAT I.  
PA (STAC/) STACHE-CRAIN B.  
PA (DICK/) DICKSON M C.  
PA (JONE/) JONES L W.  
XX  
XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
PI  
XX MPI; 2003-615964/58.  
DR  
XX New polynucleotide sequences obtained from various cDNA libraries, useful  
PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
PT mapping, in the recombinant production of protein, or in generating  
PT antisense DNA or RNA.  
XX  
XX Claim 1; SEQ ID NO 34279; 44pp; English.  
XX  
XX The invention relates to an isolated polynucleotide comprising any one of  
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
CC determined by the technique of SBH (sequencing by hybridisation). Also  
CC included is a purified polypeptide comprising a sequence corresponding to  
CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
CC are useful in diagnostics as expressed sequence tags (EST) for  
CC identifying expressed genes or for physical mapping of the human genome,  
CC in forensics, in assessing biodiversity, or in identifying mutations  
CC responsible for genetic disorders and other traits. The nucleotide  
CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
CC for chromosome and gene mapping, in the recombinant production of  
CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
CC is useful for generating antibodies specific for it. The present sequence  
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocID=20030073623  
XX

SQ Sequence 502 BP; 178 A; 97 C; 119 G; 106 T; 0 U; 2 Other;

Query Match 19.8%; Score 459; DB 9; Length 502;  
Best Local Similarity 100.0%; Pred. No. 1e-202;  
Matches 459; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 912 TTCTCTGGGAAGAACACGACGAGAGTACACAGAAAGAGCTTAACTTTAAATAAATTTGATAAGACT 971  
Dy 44 TTCTCTGGGAAGAACACGACGAGAGTACACAGAAAGAGCTTAACTTTAAATAAATTTGATAAGACT 103

Qy 972 GAACAAAAGCAAAAGTTTAAAGAAAAGAGCTTAACTTTAAATAAATTTGATAAGACT 1031  
Dy 104 GAACAAAAGCAAAAGTTTAAAGAAAAGAGCTTAACTTTAAATAAATTTGATAAGACT 163

Qy 1032 ATTAAGCTGAACCTGCTGCAGAAAAGAGCTTAACTTTAAATAAATTTGATAAGACT 1091  
Dy 164 ATTAAGCTGAACCTGCTGCAGAAAAGAGCTTAACTTTAAATAAATTTGATAAGACT 223

Qy 1092 GTGAATGCATTTAAAGAACTAGTAGCAAAATACCTCAGAGTCCACGAGCAAGATATCGG 1151  
Dy 224 GTGAATGCATTTAAAGAACTAGTAGCAAAATACCTCAGAGTCCACGAGCAAGATATCGG 283

Qy 1152 AAGCGCAGTGTGAGGATGTTGGCTGAGAAGAGAGAGTAATGAGTGTCTACGTGGA 1211  
Dy 284 AAGCGCAGTGTGAGGATGTTGGCTGAGAAGAGAGAGTAATGAGTGTCTACGTGGA 343

Qy 1212 GCCATCGAGACCTACCAAGAGTGGCCAGCCTACCTGATGTCCTGCGAGACCTGCTGAAG 1271  
Dy 344 GCCATCGAGACCTACCAAGAGTGGCCAGCCTACCTGATGTCCTGCGAGACCTGCTGAAG 403

Qy 1272 CTGAGTTTGAAGCGTCGCTCAGACAGGCAACAATTTCTAGTTCATATGAGAGGTTCCCTG 1331  
Dy 404 CTGAGTTTGAAGCGTCGCTCAGACAGGCAACAATTTCTAGTTCATATGAGAGGTTCCCTG 463

Qy 1332 CTTACCTCGCAGAGATTAGTTCAACTATTTCCCAATGAT 1370  
Dy 464 CTTACCTCGCAGAGATTAGTTCAACTATTTCCCAATGAT 502

RESULT 18  
ABQ59507/c  
ID ABQ59507 standard; cdna; 660 BP.  
AC ABQ59507;  
XX 02-AUG-2002 (first entry)  
DX Human colon cancer related nucleotide sequence SEQ ID NO:3202.  
DE Human; colon cancer; cancer; tissue profiling; forensic; mapping;  
KW genetic analysis; diagnostic; antisense therapy; gene; ss.  
XX Homo sapiens.  
OS WO200229086-A2.  
XX 11-APR-2002.  
XX 02-OCT-2001; 2001WO-0303732.  
XX 02-OCT-2000; 2000US-0237271P.  
XX (FARB ) BAYER CORP.  
XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;  
PI Thiagalingam A, Lewis ME;  
XX WPI; 2002-426115/45.  
XX New isolated nucleic acid that is differentially expressed in cancer  
PT tissues useful for determining the presence of colon cancer in a cell or  
PT tissue type, and in antisense therapy.  
XX

PS Claim 1; Fig 1; 796pp; English.  
XX ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially  
CC expressed in cancer tissues AB578993 to AB579004 represent proteins  
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be  
CC used in antisense therapy. An antibody immunoreactive with a polypeptide  
CC encoded by (I) is useful for detecting cancer in a patient sample, and  
CC for detecting the presence or absence of a polynucleotide encoded by a  
CC nucleic acid which hybridizes to (I) in a cell. A probe/primer derived  
CC from (I) can be used for determining the presence of a nucleic acid which  
CC hybridizes to (I), and for determining the phenotype of cells in a sample  
CC of cells from a patient. (I) is useful for determining the presence of  
CC colon cancer in a cell or tissue type, for determining the presence of  
CC state of other type of cancer, in antisense therapy, to generate  
CC macroarrays on a solid surface, to identify a chromosome on which the  
CC corresponding gene resides, and in tissue profiling, forensics, genetic  
CC analysis, mapping and diagnostic applications. (I) can be used to raise  
CC antibodies, and to screen for peptide analogues and antagonists  
XX

SQ Sequence 660 BP; 131 A; 181 C; 123 G; 221 T; 0 U; 4 Other;  
Query Match 15.4%; Score 359; DB 6; Length 660;  
Best Local Similarity 100.0%; Pred. No. 3.4e-156;  
Matches 359; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 581 TGATGTAGATGATAGATTTCAGAGACCCCTGGAACTTCAAGTATCTCATGAAGAAACCGACCA 640  
Dy 462 TGATGTAGATGATAGATTTCAGAGACCCCTGGAACTTCAAGTATCTCATGAAGAAACCGACCA 403

Qy 641 TAGTTACACGTGGAAGAGACAGTTTCAAGAGCTGTAAATCAGGATATGAGAGATGAT 700  
Dy 402 TAGTTACACGTGGAAGAGACAGTTTCAAGAGCTGTAAATCAGGATATGAGAGATGAT 343

Qy 701 GTCTGACGAGAAATCCAGATTCAGTGAACAGTATAGAGTAAAGATGAAAGATTGACCA 760  
Dy 342 GTCTGACGAGAAATCCAGATTCAGTGAACAGTATAGAGTAAAGATGAAAGATTGACCA 283

Qy 761 TGATACAGATGATGTAACATACCAAGTCTATGAGGAAACAAAGCAGTATATGAACCTCTAGA 820  
Dy 282 TGATACAGATGATGTAACATACCAAGTCTATGAGGAAACAAAGCAGTATATGAACCTCTAGA 223

Qy 821 AAATGAAGGATGATAAATACAGAGTAACTGCTCCCTCGAGGATATCTCTGAGAAGA 880  
Dy 222 AAATGAAGGATGATAAATACAGAGTAACTGCTCCCTCGAGGATATCTCTGAGAAGA 163

Qy 881 TTCACAGTAAATGTAAGAAGTAAGCAATTTTCTGTGGAAGAAACAGCAGGAGATAC 939  
Dy 162 TTCACAGTAAATGTAAGAAGTAAGCAATTTTCTGTGGAAGAAACAGCAGGAGATAC 104

RESULT 19  
ADT94758  
ID ADT94758 standard; cdna; 495 BP.  
XX AC ADT94758;  
XX 16-DEC-2004 (first entry)  
XX Colon cancer associated human cdna sequence #277.  
DE Colon cancer; T cell; tumour protein; C634S; C635S; C637S; C640S; C636S;  
KW humoral immune response; cellular immune response; cytostatic;  
KW immunostimulant; human; ss.  
XX Homo sapiens.  
OS US2003087818-A1.  
XX PN 08-MAY-2003.  
XX PD 01-FEB-2002; 2002US-00066543.  
XX PF 02-FEB-2001; 2001US-0267400P.  
XX PR

PR 07-FEB-2001; 2001US-0267382P.  
PR 11-MAY-2001; 2001US-0290322P.  
PR 12-JUL-2001; 2001US-0305265P.  
PR 16-AUG-2001; 2001US-0313077P.  
XX (CORI-) CORIXA CORP.  
XX  
XX Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;  
XX Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;  
XX WPI; 2003-040540/03.  
XX  
XX New isolated nucleic acids and polypeptides capable of eliciting a  
XX humoral and/or cellular immune response, useful for diagnosing,  
XX preventing or treating cancer, particularly colon cancer.  
XX  
XX Claim 1; SEQ ID NO 277; 87pp; English.  
XX  
XX The invention relates to polynucleotide and polypeptide sequences  
XX associated with cancer, particularly colon cancer. Also disclosed are (i)  
XX an expression vector comprising the polynucleotide, (ii) a host cell  
XX transformed or transfected with the expression vector, (iii) an isolated  
XX antibody, or its antigen-binding fragment, which specifically binds to  
XX the polypeptide, (iv) a method of detecting or determining the presence  
XX of cancer in a patient, (v) a fusion protein comprising at least one of  
XX the polypeptides, (vi) an oligonucleotide that hybridizes to the  
XX polynucleotide sequence under highly stringent conditions, and (vii) a  
XX method of stimulating and/or expanding T cells specific for a tumor  
XX protein. The polypeptide specifically comprises the amino acid sequence  
XX of C634S, C635S, C637S, C640S, C636S or one of the potential open reading  
XX frames (ORFs) of C636S. These polypeptides are encoded by the  
XX polynucleotide sequences, where both are capable of eliciting a humoral  
XX and/or cellular immune response. The polynucleotides, polypeptides, and  
XX antibodies are useful for diagnosing, preventing or treating cancer,  
XX particularly colon cancer. The polynucleotide and polypeptide sequences  
XX are also useful in DNA strand invasion, antisense inhibition, mutational  
XX analysis, nucleic acid purification, isolation of transcriptionally  
XX active genes, blocking or transcription factor binding, genome cleavage  
XX or in situ hybridization, and as enhancers of transcription or  
XX biomarkers. This sequence represents a human colon cancer associated  
XX cDNA. Note: The sequence data for this patent was obtained in electronic  
XX format directly from the USPTO web site at seqdata.uspto.gov  
XX  
XX Sequence 495 BP; 183 A; 95 C; 111 G; 104 T; 0 U; 2 Other;  
XX  
XX Query Match 10.7%; Score 249; DB 11; Length 495;  
XX Best Local Similarity 99.7%; Pred No. 4.9e-105;  
XX Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 630 GAACCGAGCATAGTTACCACTGGGAGAGACAGTTTCAAGACTGTATTCAGGATATG 689  
DB |||||||  
QY 41 GAACCGAGCATAGTTACCACTGGGAGAGACAGTTTCAAGACTGTATTCAGGATATG 100  
DB |||||||  
QY 690 GAAGAGATGATGTCGTAGCAGGAGAAATCCAGATTCAGTGAACACAGTAGTAGAAGATGAA 749  
DB |||||||  
QY 101 GAAGAGATGATGTCGTAGCAGGAGAAATCCAGATTCAGTGAACACAGTAGTAGAAGATGAA 160  
DB |||||||  
QY 750 AGATTGACCATGATACAGATGATGTAACTACCAAGTCTATAGGAAACAGCAGTATAT 809  
DB |||||||  
QY 161 AGATTGACCATGATACAGATGATGTAACTACCAAGTCTATAGGAAACAGCAGTATAT 220  
DB |||||||  
QY 810 GAACCTCTAGAAATGAAGGGATAGAAATCAGAGTAATCTGCTCCCTGAGGATAT 869  
DB |||||||  
QY 221 GAACCTCTAGAAATGAAGGGATAGAAATCAGAGTAATCTGCTCCCTGAGGATAT 280  
DB |||||||  
QY 870 CCTGTAGAAATTCACAGGTAATGTAGAAAGTAAGTACATTTTCTGTGGAAAGACAG 929  
DB |||||||  
QY 281 CCTGTAGAAATTCACAGGTAATGTAGAAAGTAAGTACATTTTCTGTGGAAAGACAG 340  
DB |||||||  
RESULT 20  
ACH93769/c  
ID ACH93769 standard; DNA; 161 BP.

XX ACH93769;  
XX 29-JUL-2004 (first entry)  
XX Human genome derived single exon probe #26964.  
XX Human; probe; ss; gene expression; single exon probe; microarray;  
XX alternative splicing event; genomic alteration.  
XX Homo sapiens.  
XX US2003194704-A1.  
XX 16-OCT-2003.  
XX 03-APR-2002; 2002US-00029386.  
XX 03-APR-2002; 2002US-00029386.  
XX (PENN/) PENN S G.  
XX (RANK/) RANK D R.  
XX (HANZ/) HANZEL D K.  
XX Penn SG, Rank DR, Hanzel DK;  
XX WPI; 2004-119264/12.  
XX New human genome-derived single exon nucleic acid probes useful for human  
XX gene expression analysis, for identifying or characterizing alternative  
XX splicing events, for assessing genomic alterations or as tools for  
XX surveying tissues.  
XX Claim 1; SEQ ID NO 26964; 80pp; English.  
XX The invention relates to a nucleic acid probe for measuring human gene  
XX expression, comprising any of the 27,400 fully defined nucleotide  
XX sequences in the specification, or their complements or fragments, and  
XX encoding at least 8 amino acids of any of the 6888 amino acid sequences  
XX fully defined in the specification. The probe is a single exon probe that  
XX hybridizes under high stringency conditions to a nucleic acid molecule  
XX expressed in human cells or tissues. Also included are a spatially-  
XX addressable set of single exon nucleic acid probes for measuring human  
XX gene expression (comprising a plurality of single exon nucleic acid  
XX probes cited above, where each of the plurality of probes is separately  
XX and addressably isolatable or amplifiable from the plurality), a single  
XX exon microarray for measuring human gene expression, a method of  
XX measuring human gene expression, an ORF-encoded peptide comprising at least 8  
XX probe cited above, an ORF-encoded peptide comprising at least 8  
XX contiguous amino acids of any of the above-mentioned amino acid  
XX sequences (optionally with conservative amino acid substitutions), an  
XX isolated antibody that binds specifically to a peptide cited above,  
XX a method of selling and/or licensing single exon probes or microarrays to  
XX a customer desiring to measure gene expression, a method of providing  
XX human gene expression data by subscription, and a computer-readable  
XX storage medium which contains a database having a plurality of records  
XX (each record including data on the expression of a single exon probe  
XX cited above). The probe, methods and apparatus are useful in gene  
XX expression analysis. The probes may be used as tools for surveying  
XX tissues to detect the presence of expressed messages that contain their  
XX specific exon, or in constructing genome-derived single exon microarrays.  
XX In addition, the probes are used in identifying and characterizing  
XX alternative splicing events, in detecting and characterizing gross  
XX alterations in the genomic locus that includes their exon, in assessing  
XX smaller genomic alterations, in priming the synthesis of nucleic acids,  
XX or in expressing the ORF-encoded peptide. The present sequence is a human  
XX single exon probe of the invention. Note: The sequence data for this  
XX patent did not form part of the printed specification, but was obtained  
XX in electronic format directly from USPTO at  
XX seqdata.uspto.gov/sequence.html?DocID=20030194704  
XX  
XX Sequence 161 BP; 45 A; 49 C; 25 G; 42 T; 0 U; 0 Other;

```
Query Match      4.5%; Score 104; DB 12; Length 161;
Best Local Similarity 100.0%; Pred. No. 1.4e-37;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 162 GGAACCTTCATTCTTCACGTCGTTTATGTTGATTCGATTCGTCGGCGTCTGGACATCTGTA 221
Db 110 GGAACCTTCATTCTTCACGTCGTTTATGTTGATTCGATTCGTCGGCGTCTGGACATCTGTA 51

Oy 222 GCTGTCGTTTGGTTGATCTTGTTCACCTATCATCAGGAAGTTCTAGG 265
Db 50 GCTGTCGTTTGGTTGATCTTGTTCACCTATCATCAGGAAGTTCTAGG 7

RESULT 21
ACH80069/C
ID ACH80069 standard; DNA; 592 BP.
AC ACH80069;
XX
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #13264.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
XX US2003194704-A1.
XX
XX 16-OCT-2003.
XX
XX 03-APR-2002; 2002US-00029386.
XX
XX 03-APR-2002; 2002US-00029386.
XX
XX (PENN/) PENN S G.
XX PA (RANK/) RANK D R.
XX PA (HANZ/) HANZEL D K.
XX
XX Penn SG, Rank DR, Hanzel DK;
XX
XX WPI; 2004-119264/12.
XX
XX New human genome-derived single exon nucleic acid probes useful for human
XX gene expression analysis, for identifying or characterizing alternative
XX splicing events, for assessing genomic alterations or as tools for
XX surveying tissues.
XX
XX Claim 15; SEQ ID NO 13264; 80pp; English.
XX
XX The invention relates to a nucleic acid probe for measuring human gene
XX expression, comprising any of the 27,400 fully defined nucleotide
XX sequences in the specification, or their complements or fragments, and
XX encoding at least 8 amino acids of any of the 688 amino acid sequences
XX fully defined in the specification. The probe is a single exon probe that
XX hybridises under high stringency conditions to a nucleic acid molecule
XX expressed in human cells or tissues. Also included are a spatially-
XX addressable set of single exon nucleic acid probes for measuring human
XX gene expression (comprising a plurality of single exon nucleic acid
XX probes cited above, where each of the plurality of probes is separately
XX and addressably isolatable or amplifiable from the plurality), a single
XX exon microarray for measuring human gene expression, a method of
XX measuring human gene expression, a vector comprising the single exon
XX probe cited above, an ORF-encoded peptide comprising at least 8
XX contiguous amino acids of any of the above-mentioned amino acid
XX sequences (optionally with conservative amino acid substitutions), an
XX isolated antibody that binds specifically to a peptide cited above,
XX methods of selling and/or licensing single exon probes or microarrays to
XX a customer desiring to measure gene expression, a method of providing
XX human gene expression data by subscription, and a computer-readable
XX storage medium which contains a database having a plurality of records
XX (each record including data on the expression of a single exon probe
```

```
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030194704
XX
XX SQ Sequence 592 BP; 220 A; 122 C; 79 G; 171 T; 0 U; 0 Other;
```

```
Query Match      4.5%; Score 104; DB 12; Length 592;
Best Local Similarity 100.0%; Pred. No. 1.4e-37;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 162 GGAACCTTCATTCTTCACGTCGTTTATGTTGATTCGATTCGTCGGCGTCTGGACATCTGTA 221
Db 230 GGAACCTTCATTCTTCACGTCGTTTATGTTGATTCGATTCGTCGGCGTCTGGACATCTGTA 171

Oy 222 GCTGTCGTTTGGTTGATCTTGTTCACCTATCATCAGGAAGTTCTAGG 265
Db 170 GCTGTCGTTTGGTTGATCTTGTTCACCTATCATCAGGAAGTTCTAGG 127
```

```
RESULT 22
AAAX22124
ID AAAX22124 standard; DNA; 1382 BP.
```

```
AC AAAX22124;
XX
XX 18-MAY-1999 (first entry)
XX
DE Human secreted protein gene 14 clone HAGBZ81.
```

```
XX Human; secreted protein; gene therapy; protein therapy; tissue; cancer;
XX tumour; neurodegenerative disorder; leukaemia; autoimmune disease; AIDS;
XX developmental abnormality; foetal deficiency; Alzheimer's disease;
XX cognitive disorder; schizophrenia; immunological disorder; mood disorder;
XX immune deficiency disease; respiratory disorder; arthritis; skeletal;
XX haematopoietic disorder; neural; osteoporosis; metabolic disorders;
XX cardiovascular; endocrine; gastrointestinal; asthma; diagnosis; ds.
```

```
OS Homo sapiens.
XX
XX WO9901020-A2.
```

```
XX 14-JAN-1999.
XX
XX 30-JUN-1998; 98WO-US013608.
XX
XX 01-JUL-1997; 97US-0051381P.
XX 01-JUL-1997; 97US-0051480P.
XX 12-SEP-1997; 97US-0058598P.
XX 12-SEP-1997; 97US-0058663P.
```

```
XX (HUMA-) HUMAN GENOME SCI INC.
```

```
XX Carter KC, Feng P, Rosen CA, Ruben SM, Endress GA;
```

```
XX WPI; 1999-105683/09.
```

```
XX P-PSDB; AA01148, AA01190, AA01191, AA01192, AA01193.
```

```
XX New isolated human genes and the secreted polypeptides they encode -
XX useful for diagnosis and treatment of e.g. cancers, neurological
XX disorders, immune diseases, immune deficiency diseases or blood
XX disorders.
```

```
XX Claim 4; Page 131; 179pp; English.
```







Db 293 AGTTTATTAGG 304

RESULT 26  
ADN95564/c  
ID ADN95564 standard; DNA; 60 BP.  
XX AC ADN95564;  
XX DT 01-JUL-2004 (first entry)  
XX DE Human BEC/LEC-related gene sequence SeqID487.  
XX KW growth; differentiation; blood endothelial cell; BEC;  
XX KW lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGFR-3;  
XX KW lymphatic growth agent; VEGF-C; VEGF-D; antiangiogenic; cytostatic;  
XX KW vasotropic; antiinflammatory; gene therapy; endothelial cell disorder;  
XX KW inflammatory disease; cancer metastasis; lymphatic system; gene; ds;  
XX KW human.  
XX OS Homo sapiens.  
XX PN WO200210449-A2.  
XX PD 07-FEB-2002.  
XX PF 20-JUL-2001; 2001WO-IB001903.  
XX PR 28-JUL-2000; 2000US-0221607P.  
XX PR 02-MAY-2001; 2001US-0287724P.  
XX PA (COMP-) COMPUGEN INC.  
XX PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;  
XX WIPI; 2002-257383/30.  
XX DR  
XX PT New oligonucleotide libraries comprising oligonucleotides which  
PT selectively hybridize to mRNAs transcribed from a transcription unit of a  
PT genome, useful for detecting tissue-, pathology-, and developmental-  
PT specific genes.  
XX PS Example 1; SEQ ID NO 11070; 47pp; English.  
XX CC The present invention describes oligonucleotide libraries for detecting  
CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-  
CC )transcriptome comprises messenger RNAs transcribed from multiple  
CC transcription units that populate a genome. The library comprises several  
CC oligonucleotides, each capable of hybridising selectively to a set of  
CC messenger RNAs transcribed from a given transcription unit of the genome,  
CC which encodes one or more messenger RNA splice variants. The  
CC oligonucleotide libraries are useful for detecting mRNAs from a  
CC biological sample, in expression profiling studies, in qualitatively or  
CC quantitatively characterising the corresponding transcriptome, and in  
CC detecting RNA transcripts and splice variants of human or animal  
CC transcriptomes. The libraries may also be used as specialised mini  
CC libraries to detect transcripts of a sub-transcriptome under a particular  
CC - and pathology-specific state, and so allowing the detection of tissue  
CC specific tissue under a specific pathological condition; to detect  
CC developmental specific genes; and to detect RNA transcripts and splice  
CC variants of a transcriptome of a patient suffering from a particular  
CC disorder. ABN7253 to ABN9589 represent oligonucleotide sequences from  
CC rats, humans and mice, which are used in the exemplification of the  
CC present invention. N.B. The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIFO at ftp.wifo.int/pub/published\_pct\_sequences

XX SQ Sequence 60 BP; 22 A; 8 C; 17 G; 13 T; 0 U; 0 Other;  
Query Match 2.6%; Score 60; DB 6; Length 60;  
Best Local Similarity 100.0%; Pred. No. 4.2e-17;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 672 GACTTATCAGGATATGGAAGAGATGATCTGACGAGGAATTCAGATTCAGTGA 731  
Db 1 GACTTATCAGGATATGGAAGAGATGATCTGACGAGGAATTCAGATTCAGTGA 60

RESULT 27  
ADN95564/c  
ID ADN95564 standard; DNA; 2208 BP.  
XX AC ADN95564;  
XX DT 01-JUL-2004 (first entry)  
XX DE Human BEC/LEC-related gene sequence SeqID487.  
XX KW growth; differentiation; blood endothelial cell; BEC;  
XX KW lymphatic endothelial cell; LEC; hereditary lymphoedema; VEGFR-3;  
XX KW lymphatic growth agent; VEGF-C; VEGF-D; antiangiogenic; cytostatic;  
XX KW vasotropic; antiinflammatory; gene therapy; endothelial cell disorder;  
XX KW inflammatory disease; cancer metastasis; lymphatic system; gene; ds;  
XX KW human.  
XX OS Homo sapiens.  
XX PN WO2003080640-A1.  
XX PD 02-OCT-2003.  
XX PF 07-MAR-2003; 2003WO-US006900.  
XX PR 07-MAR-2002; 2002US-0363019P.  
XX PA (LUDW-) LUDWIG INST CANCER RES.  
XX PA (LICN) LICENTIA LTD.  
XX PI Alitalo K, Makinen T, Petrova T, Saharinen P, Saharinen J;  
XX WIPI; 2003-876899/81.  
XX DR P-PSDB; ADN95563.  
XX PS Example 1; SEQ ID NO 487; 176pp; English.  
XX CC This invention relates to a method of differentially modulating the  
CC growth or differentiation of blood endothelial cells (BEC) or lymphatic  
CC endothelial cells (LEC) comprises contacting endothelial cells with a  
CC composition comprising an agent that differentially modulates blood or  
CC lymphatic endothelial cells. Treating hereditary lymphoedema comprises  
CC identifying a human subject with lymphoedema and with a mutation in at  
CC least one allele of a gene encoding a LEC protein, where the mutation  
CC correlates with lymphoedema in human subjects, and with the proviso that  
CC the LEC protein is not VEGFR-3; and administering to the subject a  
CC composition comprising a lymphatic growth agent selected from VEGF-C or  
CC VEGF-D polypeptides and polynucleotides. The invention may be useful for  
CC the development of compounds with an antiangiogenic, cytostatic,  
CC vasotropic or antiinflammatory activity or for gene therapy. The method  
CC is useful in modulating the growth or differentiation of blood  
CC endothelial cells or lymphatic endothelial cells; in treating hereditary  
CC lymphoedema; in screening for an endothelial cell disorder or  
CC predisposition to the disorder or in monitoring the efficacy or toxicity  
CC of a drug on endothelial cells. The agent is useful in manufacturing a  
CC medicament for the differential modulation of blood vessel endothelial  
CC cell or lymphatic vessel endothelial cell growth or differentiation. The  
CC lymphatic growth agent may also be used in manufacturing a medicament for  
CC the treatment of hereditary lymphoedema resulting from a mutation in a  
CC LEC gene or of other diseases involving the lymphatic vessels, such as  
CC various inflammatory diseases and cancer metastasis via the lymphatic  
CC system. The present sequence is that of a human LEC/BEC differentially  
CC expressed gene which is related to the method of the invention. Note: This  
CC sequence does not appear in the specification but was obtained by the  
CC indexer using the source data given in table 14 of the specification.  
XX SQ Sequence 2208 BP; 621 A; 428 C; 484 G; 675 T; 0 U; 0 Other;

Query Match 2.5%; Score 59; DB 11; Length 2208;  
Best Local Similarity 100.0%; Pred. No. 1.1e-16;  
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2239 GGCATCGGAAGTACACACGAGCAGACGAGCCTTCCAGCAATTTAGCATGAATTC 2297  
 DB 59 GGCATCGGAAGTACACACGAGCAGACGAGCCTTCCAGCAATTTAGCATGAATTC 1

RESULT 28  
 ADRI14176/C  
 ID ADRI14176 standard; DNA; 2208 BP.  
 XX AC  
 XX ADRI14176;  
 XX  
 DT 21-OCT-2004 (first entry)  
 XX  
 DE Human NF-kappaB pathway-associated gene SeqID177.  
 XX  
 XX NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;  
 KW antiarthritic; antirheumatic; gastrointestinal-Gen; antiasthmatic;  
 KW antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;  
 KW immunosuppressive; vulnary; gene therapy; immune disorder;  
 KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;  
 KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;  
 KW hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;  
 KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;  
 KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;  
 KW viral replication; host cell survival; evasion of immune response;  
 KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;  
 KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;  
 KW autoimmune disorder; hyper immune activity;  
 KW aberrant acute phase response; hypercongenital condition; birth defect;  
 KW necrotic lesion; wound; organ transplant rejection;  
 KW aberrant signal transduction; proliferating disorder; cancer;  
 KW HIV propagation; gene; ds; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004065577-A2.  
 XX  
 PD 05-AUG-2004.  
 XX  
 PF 13-JAN-2004; 2004WO-US000798.  
 XX  
 XX 14-JAN-2003; 2003US-0440068P.  
 PR 12-MAY-2003; 2003US-0469757P.  
 XX  
 PA (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 XX  
 XX Nadler SG, Neubauer MG, Feder JN, Carman J;  
 PI P-PSDB; ADRI14177.  
 XX  
 DR New isolated polynucleotides and polypeptides associated with NF-kappaB  
 PT pathway, useful for diagnosing, treating, or preventing disorders or  
 PT diseases associated with NF-kappaB pathway.  
 XX  
 PS Claim 1; SEQ ID NO 177; 237pp; English.  
 XX  
 PS This invention relates to the novel association of protein sequences (and  
 CC the genes which encode them) to the NF-kappaB pathway. The invention may  
 CC be useful for the production of compounds with an antiinflammatory,  
 CC cytostatic, hepatotropic, virucide, antiarthritic, antirheumatic,  
 CC gastrointestinal-Gen, antiasthmatic, antiarteriosclerotic,  
 CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or  
 CC vulnary activity or for gene therapy. The proteins and nucleotides are  
 CC useful for diagnosing, preventing, treating, or ameliorating conditions  
 CC or diseases associated with the NF-kappaB pathway. The condition is an  
 CC immune disorder, an inflammatory disorder, an inflammatory disorder  
 CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,  
 CC hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM  
 CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic  
 CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,  
 CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell

CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory  
 CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick  
 CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper  
 CC immune activity, disorders related to aberrant acute phase responses,  
 CC hypercongenital conditions, birth defects, necrotic lesions, wounds,  
 CC organ transplant rejection, conditions related to organ transplant  
 CC rejection, disorders related to aberrant signal transduction, infected  
 CC proliferating disorders, cancers and HIV propagation in cells which is  
 CC subject to the novel association with the NF-kappaB pathway of the  
 CC invention. Note: This sequence does not appear in the specification but  
 CC was obtained by the indexer from Genbank.

XX SQ Sequence 2208 BP; 621 A; 428 C; 484 G; 675 T; 0 U; 0 Other;

QY Query Match 2.5%; Score 59; DB 13; Length 2208;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-16;  
 Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2239 GGCATCGGAAGTACACACGAGCAGACGAGCCTTCCAGCAATTTAGCATGAATTC 2297  
 DB 59 GGCATCGGAAGTACACACGAGCAGACGAGCCTTCCAGCAATTTAGCATGAATTC 1

## RESULT 29

ADP23153/C

ID ADP23153 standard; cDNA; 2208 BP.

XX AC

ADP23153;

XX

DT 18-NOV-2004 (first entry)

DE

PRO polypeptide encoding cDNA SEQ ID NO:247.

XX

ss; gene; PRO; antiinflammatory; antiarthritic; antirheumatic;  
 KW immunosuppressive; osteopathic; antidiabetic; dermatological;  
 KW antipsoriatic; antiallergic; antiasthmatic; hepatotropic; respiratory;  
 KW gene therapy; immune system.

XX Unidentified.

OS

WO2004041170-A2.

PN

21-MAY-2004.

PD

30-OCT-2003; 2003WO-US034312.

PF

01-NOV-2002; 2002US-0423394P.

PR

(GETH ) GENENTECH INC.

PA

Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;

PI Wu TD;

XX

WPI; 2004-419628/39.

DR

P-PSDB; ADP23154.

XX

New PRO polypeptides and polynucleotides, useful for treating e.g.  
 PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated  
 PT renal disease, or demyelinating diseases of the central or peripheral  
 PT nervous system.

XX Claim 1; SEQ ID NO 247; 2940pp; English.

PS

The invention relates to a novel isolated nucleic acid and the PRO  
 CC polypeptide encoded by it. A protein of the invention has  
 CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,  
 CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,  
 CC antiasthmatic, hepatotropic, and respiratory activity. A polynucleotide  
 CC of the invention may have a use in gene therapy. The PRO polypeptide, its  
 CC agonist, antagonist, or antibody that specifically binds to the  
 CC polypeptide is useful for treating an immune related disorder such as  
 CC systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,

CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an  
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic  
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune  
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal  
CC disease, a demyelinating disease of the central or peripheral nervous  
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,  
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary  
CC disease, infectious or autoimmune chronic active hepatitis, primary  
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,  
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's  
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin  
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic  
CC disease, asthma, allergic rhinitis, atopic dermatitis, food  
CC hypersensitivity, urticaria, an immunologic disease of the lung,  
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity  
CC pneumonitis, a transplantation associated disease, graft rejection or  
CC graft-versus-host disease. The present sequence encodes a PRO protein of  
CC the invention.

XX Sequence 2208 BP; 621 A; 428 C; 484 G; 675 T; 0 U; 0 Other;

Query Match 2.5%; Score 59; DB 13; Length 2208;  
Best Local Similarity 100.0%; Pred. No. 1.1e-16;  
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2239 GGCATCCGCACTGACACACAGAGAGCGAGCGCTTCCAGCAATTTAGCATGATTC 2297  
DB 59 GGCATCCGCACTGACACACAGAGAGCGAGCGCTTCCAGCAATTTAGCATGATTC 1

RESULT 30

ABQ59571/c  
ID ABQ59571 standard; cDNA; 583 BP.  
XX AC ABQ59571;  
XX XX  
XX 02-AUG-2002 (first entry)  
XX Human colon cancer related nucleotide sequence SEQ ID NO:3266.  
XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;  
XX genetic analysis; diagnostic; antisense therapy; gene; ss.  
XX Homo sapiens.  
XX WO200229086-A2.  
XX 11-APR-2002.  
XX 02-OCT-2001; 2001WO-US030732.  
XX 02-OCT-2000; 2000US-0237271P.  
XX (FARB ) BAYER CORP.  
XX Burgess C, Aetle JH, Carroll E, Catino TU, Dwivedi P, Molino GA;  
XX Thiagalingam A, Lewis ME;  
XX WPI; 2002-426115/45.

XX New isolated nucleic acid that is differentially expressed in cancer  
XX tissues useful for determining the presence of colon cancer in a cell or  
XX tissue type, and in antisense therapy.

XX Claim 1; Fig 1; 796pp; English.

XX ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially  
XX expressed in cancer tissues. ABQ5993 to ABQ59004 represent proteins  
XX encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be  
XX used in antisense therapy. An antibody immunoreactive with a polypeptide  
XX encoded by (I) is useful for detecting cancer in a patient sample, and  
XX for detecting the presence or absence of a polynucleotide encoded by a  
XX nucleic acid which hybridises to (I) in a cell. A probe/primer derived

CC from (I) can be used for determining the presence of a nucleic acid which  
CC hybridises to (I), and for determining the phenotype of cells in a sample  
CC of cells from a patient. (I) is useful for determining the presence of  
CC colon cancer in a cell or tissue type, for determining the presence of  
CC state of other type of cancer, in antisense therapy, to generate  
CC macroarrays on a solid surface, to identify a chromosome on which the  
CC corresponding gene resides, and in tissue profiling, forensics, genetic  
CC analysis, mapping and diagnostic applications. (I) can be used to raise  
XX antibodies, and to screen for peptide analogues and antagonists

XX Sequence 583 BP; 170 A; 128 C; 126 G; 152 T; 0 U; 7 Other;

Query Match 2.2%; Score 51; DB 6; Length 583;  
Best Local Similarity 100.0%; Pred. No. 6e-13;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2274 CTTCCAGCAATTTAGCATGATTCATGCAAGCTTGGAACTCTGAGAGA 2324  
DB 439 CTTCCAGCAATTTAGCATGATTCATGCAAGCTTGGAACTCTGAGAGA 389

RESULT 31  
ADR97461/c  
ID ADR97461 standard; DNA; 27 BP.

XX AC ADR97461;

XX XX 02-DEC-2004 (first entry)

XX Reverse quantitative PCR primer to amplify human CDIPT DNA.

XX ss; apoptosis; cancer; inflammation; autoimmune;  
XX neurodegenerative disorder; cytostatic; antiinflammatory;  
XX immunosuppressive; neuroprotective; gene therapy; PCR; primer;  
XX real-time PCR; RT-PCR; CDIPT.

XX Homo sapiens.

XX WO2004078793-A2.

XX 16-SEP-2004.

XX 05-MAR-2004; 2004WO-GB000957.

XX 07-MAR-2003; 2003GB-00005267.

XX (EIRX-) EIRX THERAPEUTICS LTD.

XX Murphy FJ, Sheehan DE, Keating KE, Hayes I, Seera L;

XX WPI; 2004-662402/64.

XX Identifying an agent that modulates the function of an apoptosis-  
XX associated polypeptide, useful for diagnosing or treating e.g. cancer,  
XX comprises comparing the binding of the polypeptide to the candidate agent  
XX and to a control agent.

XX Example 5; Page 289; 304pp; English.

XX This invention relates to novel agents that modulates the function of  
XX human apoptosis-associated proteins specified within the specification.  
XX Specifically, it refers to a method for the identification of target  
XX genes whose expression is correlated with an early stage in the  
XX regulation of apoptosis. The present invention describes a method of  
XX contacting either candidate agents or control agents to the target genes  
XX and assessing the difference of binding and inhibitory activity, where  
XX the candidate agent is selected from a low molecular weight organic  
XX molecule, an antibody or fragment thereof, an antisense oligonucleotide,  
XX a small inhibitory dsRNA, or a ribozyme. As such, the compositions and  
XX methods are useful for diagnosing and treating diseases or conditions  
XX associated with abnormal apoptosis in mammalian tissue, such as cancer,  
XX inflammation, autoimmune or neurodegenerative disorders. Accordingly,  
XX they exhibit cytostatic, antiinflammatory, immunosuppressive and

CC neuroprotective activities. These may also be used for drug screening  
CC purposes and in gene therapy. This oligonucleotide sequence is a  
CC quantitative PCR primer used to amplify a target gene associated with  
CC apoptosis regulation, given in an exemplification of the invention.  
XX  
SQ Sequence 27 BP; 10 A; 11 C; 1 G; 5 T; 0 U; 0 Other;  
Query Match 1.2%; Score 27; DB 13; Length 27;  
Best Local Similarity 100.0%; Pred. No. 0.096;  
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 286 CTGATGGTGGAGATTTCGATGCG 312  
DB 27 CTGATGGTGGAGATTTCGATGCG 1  
RESULT 32  
ADR97432  
ID ADR97432 standard; DNA; 23 BP.  
XX  
AC ADR97432;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
DE Forward quantitative PCR primer to amplify human Cathepsin C1 DNA.  
XX  
KW ss; apoptosis; cancer; inflammation; autoimmune;  
KW neurodegenerative disorder; cytostatic; antiinflammatory;  
KW immunosuppressive; neuroprotective; gene therapy; PCR; primer;  
KW real-time PCR; RT-PCR; Cathepsin C1.  
XX  
OS Homo sapiens.  
XX  
PN WO2004078783-A2.  
XX  
PD 16-SEP-2004.  
XX  
PF 05-MAR-2004; 2004WO-GB000957.  
XX  
PR 07-MAR-2003; 2003GB-00005267.  
XX  
PA (EIRX-) EIRX THERAPEUTICS LTD.  
XX  
PI Murphy FJ, Sheehan DE, Keating KE, Hayes I, Seera L;  
XX WPI; 2004-662402/64.  
XX  
PT Identifying an agent that modulates the function of an apoptosis-  
PT associated polypeptide, useful for diagnosing or treating e.g. cancer,  
PT comprises comparing the binding of the polypeptide to the candidate agent  
PT and to a control agent.  
XX  
PS Example 5; Page 289; 304pp; English.  
XX  
CC This invention relates to novel agents that modulates the function of  
CC human apoptosis-associated proteins specified within the specification.  
CC Specifically, it refers to a method for the identification of target  
CC genes whose expression is correlated with an early stage in the  
CC regulation of apoptosis. The present invention describes a method of  
CC contacting either candidate agents or control agents to the target genes  
CC and assessing the difference of binding and inhibitory activity, where  
CC the candidate agent is selected from a low molecular weight organic  
CC molecule, an antibody or fragment thereof, an antisense oligonucleotide,  
CC a small inhibitory dRNA, or a ribozyme. As such, the compositions and  
CC methods are useful for diagnosing and treating diseases or conditions  
CC associated with abnormal apoptosis in mammalian tissue, such as cancer,  
CC inflammation, autoimmune or neurodegenerative disorders. Accordingly,  
CC they exhibit cytostatic, antiinflammatory, immunosuppressive and  
CC neuroprotective activities. These may also be used for drug screening  
CC purposes and in gene therapy. This oligonucleotide sequence is a  
CC quantitative PCR primer used to amplify a target gene associated with  
CC apoptosis regulation, given in an exemplification of the invention.  
XX

SQ Sequence 23 BP; 9 A; 5 C; 8 G; 1 T; 0 U; 0 Other;  
Query Match 1.0%; Score 23; DB 13; Length 23;  
Best Local Similarity 100.0%; Pred. No. 7;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 106 CCCGAGAGAGACAAAGCATGGA 128  
DB 1 CCCGAGAGAGACAAAGCATGGA 23  
RESULT 33  
ABL25903  
ID ABL25903 standard; DNA; 855 BP.  
XX  
AC ABL25903;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 29182.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ds.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW,  
XX WPI; 2001-656860/75.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
XX  
PS Claim 1; SEQ ID NO 29182; 2lpp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
CC AB872072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 855 BP; 215 A; 293 C; 220 G; 127 T; 0 U; 0 Other;  
Query Match 0.9%; Score 22; DB 4; Length 855;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 36 AGCAGCGGCAACAGCAGCAGCA 57  
DB 139 AGCAGCGGCAACAGCAGCAGCA 160  
RESULT 34  
ABL19239  
ID ABL19239 standard; DNA; 1933 BP.  
XX



```

Query Match          0.9%; Score 22; DB 4; Length 3705;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      36  AGCAGCGGCAACAGCAGCAGCA  57
         |||||
DB      276 AGCAGCGGCAACAGCAGCAGCA  297

RESULT 38
ABL19238
ID  ABL19238 standard; DNA; 3933 BP.
XX  ABL19238;
XX
DT  26-MAR-2002 (first entry)
XX
DE  Drosophila melanogaster genomic polynucleotide SEQ ID NO 9187.
XX
KW  Drosophila; developmental biology; cell signalling; insecticide;
   pharmaceutical; gene; ds.
XX
OS  Drosophila melanogaster.
XX
SN  WO200171042-A2.
XX
PD  27-SEP-2001.
XX
PF  23-MAR-2001; 2001WO-US009231.
XX
PR  23-MAR-2000; 2000US-0191637P.
PR  11-JUL-2000; 2000US-00614150.
XX
PA  (PEKE ) PE CORP NY.
XX
PI  Venter JC, Adams M, Li PWD, Myers EW;
   WPI; 2001-656860/75.
XX
DR  New isolated nucleic acid detection reagent for detecting 1000 or more
   genes from Drosophila and for elucidating cell signaling and cell-cell
   interactions.
XX
PS  Claim 1; SEQ ID NO 9187; 21pp + Sequence Listing; English.
XX
CC  The invention relates to an isolated nucleic acid detection reagent
   capable of detecting 1000 or more genes from Drosophila. The invention is
   useful in developmental biology and in elucidating cell signalling and
   cell-cell interactions in higher eukaryotes for the development of
   insecticides, therapeutics and pharmaceutical drugs. The invention
   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
   sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
   ABB72072). The sequence data for this patent did not form part of the
   printed specification, but was obtained in electronic format directly
   from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ  Sequence 3933 BP; 1234 A; 995 C; 781 G; 923 T; 0 U; 0 Other;

Query Match          0.9%; Score 22; DB 4; Length 3933;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      40  GCGGCAACAGCAGCAGCAGCGG  61
         |||||
DB      1484 GCGGCACAGCAGCAGCAGCGG  1505

RESULT 39
ABL09456
ID  ABL09456 standard; cDNA; 3974 BP.
XX
AC  ABL09456;

```



XX DT 26-MAR-2002 (first entry)  
 XX DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 22850.  
 XX KW Drosophila; developmental biology; cell signalling; insecticide;  
 XX KW pharmaceutical; gene; ss.  
 XX OS Drosophila melanogaster.  
 XX PN WO200171042-A2.  
 XX PD 27-SEP-2001.  
 XX PF 23-MAR-2001; 2001WO-US009231.  
 XX PR 23-MAR-2000; 2000US-0191637P.  
 XX PR 11-JUL-2000; 2000US-00614150.  
 XX PA (PEKE ) PE CORP NY.  
 XX PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX PI WPI; 2001-656860/75.  
 XX DR P-PSDB; AB865353.  
 XX DR New isolated nucleic acid detection reagent for detecting 1000 or more  
 XX PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 XX PT interactions.  
 XX PS Claim 1; SEQ ID NO 22850; 21pp + Sequence Listing; English.  
 XX CC The invention relates to an isolated nucleic acid detection reagent  
 XX CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 XX CC useful in developmental biology and in elucidating cell signalling and  
 XX CC cell-cell interactions in higher eukaryotes for the development of  
 XX CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 XX CC sequences (ABL01840-ABL16175) and the encoded proteins (AB857737-  
 XX CC AB872072). The sequence data for this patent did not form part of the  
 XX CC printed specification, but was obtained in electronic format directly  
 XX CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 3974 BP; 1173 A; 1026 C; 915 G; 860 T; 0 U; 0 Other;  
 XX  
 Query Match 0.9%; Score 22; DB 4; Length 3974;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 38 CAGCGGCAACAGCAGCAGCAGC 59  
 DB 2201 CAGCGGCAACAGCAGCAGCAGC 2222  
 RESULT 40  
 ABL23386  
 ID ABL23386 standard; DNA; 5903 BP.  
 XX AC ABL23386;  
 XX DT 26-MAR-2002 (first entry)  
 XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 21631.  
 XX KW Drosophila; developmental biology; cell signalling; insecticide;  
 XX KW pharmaceutical; gene; ds.  
 XX OS Drosophila melanogaster.  
 XX PN WO200171042-A2.  
 XX PD 27-SEP-2001.  
 XX

PF 23-MAR-2001; 2001WO-US009231.  
 XX 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX (PEKE ) PE CORP NY.  
 XX PA Venter JC, Adams M, Li PWD, Myers EW;  
 XX PI WPI; 2001-656860/75.  
 XX DR New isolated nucleic acid detection reagent for detecting 1000 or more  
 XX PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 XX PT interactions.  
 XX PS Claim 1; SEQ ID NO 21631; 21pp + Sequence Listing; English.  
 XX CC The invention relates to an isolated nucleic acid detection reagent  
 XX CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 XX CC useful in developmental biology and in elucidating cell signalling and  
 XX CC cell-cell interactions in higher eukaryotes for the development of  
 XX CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 XX CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 XX CC sequences (ABL01840-ABL16175) and the encoded proteins (AB857737-  
 XX CC AB872072). The sequence data for this patent did not form part of the  
 XX CC printed specification, but was obtained in electronic format directly  
 XX CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 5903 BP; 1547 A; 1491 C; 1506 G; 1359 T; 0 U; 0 Other;  
 XX  
 Query Match 0.9%; Score 22; DB 4; Length 5903;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 36 AGCAGCGGCAACAGCAGCAGCAGCA 57  
 DB 1276 AGCAGCGGCAACAGCAGCAGCA 1297  
 RESULT 41  
 ADR97567/c  
 ID ADR97567 standard; RNA; 21 BP.  
 XX AC ADR97567;  
 XX DT 02-DEC-2004 (first entry)  
 XX DE Antisense GMPR2 siRNA oligomer to knockout apoptosis related gene.  
 XX KW ss; DNA/RNA hybrid; apoptosis; cancer; inflammation; autoimmune;  
 XX KW neurodegenerative disorder; cytostatic; antiinflammatory;  
 XX KW immunosuppressive; neuroprotective; gene therapy; siRNA; GMPR2.  
 XX OS Homo sapiens.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 XX FT misc\_feature 20..21  
 XX FT /tag= a  
 XX FT /note= "Oligo dT bases added to the 3' end to stabilise  
 XX FT binding"  
 XX PN WO2004078783-A2.  
 XX PD 16-SEP-2004.  
 XX PF 05-MAR-2004; 2004WO-GB000957.  
 XX PR 07-MAR-2003; 2003GB-00005267.  
 XX PA (EIRX-) EIRX THERAPEUTICS LTD.  
 XX PI Murphy FU, Sheehan DS, Keating KE, Hayes I, Seera L;

XX WPI; 2004-662402/64.

XX Identifying an agent that modulates the function of an apoptosis-  
 XX associated polypeptide, useful for diagnosing or treating e.g. cancer,  
 PT comprises comparing the binding of the polypeptide to the candidate agent  
 PT and to a control agent.

XX Example 5; Page 293; 304pp; English.

XX This invention relates to novel agents that modulates the function of  
 XX human apoptosis-associated proteins specified within the specification.  
 CC Specifically, it refers to a method for the identification of target  
 CC genes whose expression is correlated with an early stage in the  
 CC regulation of apoptosis. The present invention describes a method of  
 CC contacting either candidate agents or control agents to the target genes  
 CC and assessing the difference of binding and inhibitory activity, where  
 CC the candidate agent is selected from a low molecular weight organic  
 CC molecule, an antibody or fragment thereof, an antisense oligonucleotide,  
 CC a small inhibitory dsRNA, or a ribozyme. As such, the compositions and  
 CC methods are useful for diagnosing and treating diseases or conditions  
 CC associated with abnormal apoptosis in mammalian tissue, such as cancer,  
 CC inflammation, autoimmune or neurodegenerative disorders. Accordingly,  
 CC they exhibit cytostatic, antiinflammatory, immunosuppressive and  
 CC neuroprotective activities. These may also be used for drug screening  
 CC purposes and in gene therapy. This oligonucleotide sequence is an siRNA  
 CC oligo used to knockout expression of a human gene identified to be  
 CC associated with the regulation of apoptosis, given in an exemplification  
 CC of the invention.

XX Sequence 21 BP; 4 A; 5 C; 3 G; 2 T; 7 U; 0 Other;

Query Match 0.9%; Score 21; DB 13; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 60;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1787 AACTGGAGTTTAATCCGAGA 1807  
 DB |||||  
 21 AACTGGAGTTTAATCCGAGA 1

RESULT 42  
 AAZ98400  
 ID AAZ98400 standard; DNA; 1845 BP.  
 XX  
 AC AAZ98400;

XX 19-JUN-2000 (first entry)  
 DT  
 XX Canine betal-adrenoceptor polynucleotide sequence.

XX Betal-adrenoceptor; antisense; transcription; hypertension; canine;  
 KW myocardial ischemia; hypotensive; vasotropic; ds.  
 XX  
 OS Canis familiaris.  
 XX WO200015783-A2.  
 PN  
 XX 23-MAR-2000.  
 PD  
 XX 14-SEP-1999; 99WO-US021007.  
 PF  
 XX 14-SEP-1998; 98US-00152717.  
 PR  
 XX (UYFL ) UNIV FLORIDA.  
 PA  
 XX Phillips MI, Zhang Y;  
 PI  
 XX WPI; 2000-271424/23.

XX New antisense oligonucleotides, useful for treatment and prevention of  
 PT hypertension and myocardial ischemia, are directed against mRNA encoding  
 PT the betal-adrenoceptor.

XX Example 4; Page 90-91; 144pp; English.

XX The invention provides a composition comprising at least one  
 CC oligonucleotide, of 9-35 bases, that binds specifically to part of the  
 CC mRNA, expressed from the gene encoding a mammalian betal-adrenoceptor  
 CC polypeptide. Binding of the oligo to the betal-adrenoceptor mRNA alters  
 CC transcription of the mRNA in cells. The compositions are used to treat  
 CC diseases associated with increased number or activity of the betal-  
 CC adrenoceptor polypeptide, specifically hypertension and myocardial  
 CC ischemia. The compositions can also be used to produce recombinant  
 CC (viral) vectors for use in vaccines, for treating or preventing the  
 CC diseases. The antisense oligos may also be used to detect the mRNA or  
 CC related DNA, including visualization within a cell, to generate  
 CC transgenic animals with altered betal-adrenoceptor activity, and to  
 CC screen patients for susceptibility to hypertension, by detection of  
 CC particular alleles of betal-adrenoceptor genes. The present sequence  
 CC represents the polynucleotide sequence of the canine betal-adrenoceptor  
 CC (GenBank Accn No: U73207)

XX Sequence 1845 BP; 197 A; 754 C; 647 G; 246 T; 0 U; 1 Other;

Query Match 0.9%; Score 21; DB 3; Length 1845;  
 Best Local Similarity 100.0%; Pred. No. 53;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 CAGCAGCGGCTCCGCGAGCGG 73  
 DB |||||  
 93 CAGCAGCGGCTCCGCGAGCGG 113

RESULT 43  
 ABK40732  
 ID ABK40732 standard; DNA; 1845 BP.  
 XX  
 AC ABK40732;

XX 21-MAY-2002 (first entry)  
 DT  
 XX Dog betal-adrenoceptor gene.

XX ds; antisense; betal adrenoceptor; betal-AR; vasotropic; hypotensive;  
 KW cardiant; hypertension; hypertrophy; cardiac ischaemia;  
 KW cardiovascular disease; cardiac dysfunction; gene.  
 XX  
 OS Canis familiaris.  
 XX WO200204623-A2.  
 PN  
 XX 17-JAN-2002.  
 PD  
 XX 11-JUL-2001; 2001WO-US021759.  
 PF  
 XX 11-JUL-2000; 2000US-00614034.  
 PR  
 XX (UYFL ) UNIV FLORIDA.  
 PA  
 XX Phillips MI, Zhang Y;  
 PI  
 XX WPI; 2002-164644/21.  
 DR  
 XX Novel antisense oligonucleotides that specifically bind to mRNA encoding  
 PT beta 1-adrenoceptor polypeptide, useful for treating cardiac  
 PT dysfunction, hypertension, hypertrophy and other cardiovascular diseases  
 PT in humans.  
 XX  
 PS Example 8; Page 84-85; 186pp; English.

XX The invention relates to an isolated antisense oligonucleotide of 9-35  
 CC nucleotides in length, which specifically binds to a portion of an mRNA  
 CC expressed from a gene encoding a mammalian betal-adrenoceptor (AR)  
 CC polypeptide and alters the translation of the mRNA into the betal-AR  
 CC polypeptide in a host cell expressing the mRNA. Also included are a

CC recombinant vector comprising the antisense oligonucleotide, and a host  
CC cell comprising the vector. A composition comprising the antisense  
CC oligonucleotides is useful in the manufacture of a medicament for use in  
CC treating or ameliorating hypertension, hypertrophy and cardiac ischaemia  
CC in a mammal. A composition comprising the antisense oligonucleotides is  
CC also useful for reducing the level of beta-AR polypeptide, the  
CC transcription of beta-AR polypeptide-specific mRNA in a mammalian host  
CC cell, particularly human cell, and for decreasing blood pressure in a  
CC mammal, where the antisense oligonucleotide is operably linked to a  
CC promoter capable of expressing the oligonucleotide in the cell. A  
CC composition comprising a selected nucleic acid segment that comprises a  
CC full-length, or is a full length beta1-adrenoceptor antisense  
CC polynucleotide operatively linked to a promoter capable of expressing the  
CC polynucleotide in a cell is also useful for reducing the level of beta1-  
CC adrenoceptor polypeptide in a mammalian host cell. The antisense  
CC oligonucleotide is also useful for other cardiovascular diseases and  
CC cardiac dysfunction in humans. The present sequence is a mammalian beta-  
CC AR gene, the mRNA for which is a target for antisense oligonucleotides of  
CC the invention

SQ Sequence 1845 BP; 197 A; 754 C; 647 G; 246 T; 0 U; 1 Other;

Query Match 0.9%; Score 21; DB 6; Length 1845;  
Best Local Similarity 100.0%; Pred. NO. 53;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 CAGCAGCGGCTCCGACGCG 73  
DB 93 CAGCAGCGGCTCCGACGCG 113

RESULT 44

AAAX60124  
ID AAX60124 standard; CDNA; 4011 BP.

AC AAX60124;

DT 05-AUG-1999 (first entry)

DE DNA sequence of the 5' region/exon 1/intron 1 of murine SAP gene.

XX Signalling lymphocyte activation molecule; SLAM; SLAM associated protein;  
XX SAP; EAT-2; antigen-specific T cell activation;  
XX X-linked proliferative disease; XLP; tuberculosis; AIDS; schistosomiasis;  
XX cancer; breast cancer; prostate cancer; leukemia; leprosy;  
XX autoimmune disease; ss.

XX Mus musculus.

XX WO926980-A1.

XX 03-JUN-1999.

XX 19-NOV-1998; 98WO-US024976.

XX 21-NOV-1997; 97US-00976096.

XX 04-SEP-1998; 98US-0099160P.

XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

XX Terhorst CP, Sayos-Ortega J;

XX WPI; 1999-357816/30.

XX New signaling lymphocyte activation molecule associated protein useful in  
XX the treatment of tuberculosis and AIDS.

XX Disclosure; Fig 15; 124pp; English.

XX The present sequence represents part of a signalling lymphocyte  
XX activation molecule (SLAM) associated protein (SAP) gene. A SAP family  
XX member polypeptide (such as EAT-2) or a fragment, mutant or fusion can be  
XX used for treating a disease involving aberrant SH2 domain-containing

CC protein-mediated signal transduction in a patient. Upregulation of SAP  
CC family members can be used for the treatment of diseases involving  
CC insufficient antigen-specific T cell activation (e.g. X-linked  
CC proliferative disease (XLP), tuberculosis, AIDS, schistosomiasis) or  
CC cancer (e.g. breast cancer, prostate cancer, leukemia), or leprosy.  
CC Downregulation of SAP family members can be used for the treatment of  
CC diseases characterized by an excessive amount of antigen-specific T cell  
CC activation (e.g. autoimmune diseases). The products can also be used for  
CC detection and diagnosis

SQ Sequence 4011 BP; 1125 A; 828 C; 885 G; 1173 T; 0 U; 0 Other;

Query Match 0.9%; Score 21; DB 2; Length 4011;  
Best Local Similarity 100.0%; Pred. NO. 51;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 GCAGCGGCACAGCAGCAGCA 57  
DB 3157 GCAGCGGCACAGCAGCAGCA 3177

RESULT 45

AAF89814/c

ID AAF89814 standard; DNA; 20 BP.

AC AAF89814;

DT 23-JUL-2001 (first entry)

DE Human aspartyl (asparaginyl) beta-hydroxylase antisense molecule.

XX Epidermal growth factor-like domain; EGF-like domain; cancer;

XX human aspartyl beta-hydroxylase; HAAH; malignant neoplasm; tumour;

XX antisense oligonucleotide; ss.

XX Homo sapiens.

XX WO200135102-A2.

XX 17-MAY-2001.

XX 08-NOV-2000; 2000WO-US030738.

XX 08-NOV-1999; 99US-00436184.

XX (RHOD-) RHODE ISLAND HOSPITAL LIFESPAN PARTNER.

XX Wands JR, De La Monte SM, Ince N, Carlson RI;

XX WPI; 2001-329171/34.

XX Diagnosing malignant neoplasm in a mammal, involves contacting mammalian  
XX sample with antibody that binds to human aspartyl beta-hydroxylase  
XX polypeptide to form antigen-antibody complex and detecting the complex.

XX Claim 16; Page 57; 76pp; English.

XX The present sequence represents an antisense oligonucleotide which  
XX inhibits human aspartyl (asparaginyl) beta-hydroxylase (HAAH) gene.  
XX Epidermal growth factor (EGF)-like domains of polypeptides are  
XX hydroxylated by HAAH enzymes. HAAH is used in the method of the  
XX invention. The specification describes a method for diagnosing a  
XX malignant neoplasm in a mammal. The method comprises contacting a body  
XX fluid with an antibody which binds to HAAH polypeptide under complex  
XX forming conditions, and detecting the antigen-antibody complex. The  
XX method is useful for diagnosing and prognosing a malignant neoplasm in a  
XX bodily fluid e.g. central nervous system (CNS)-derived body fluid, blood,  
XX serum, urine, saliva, sputum, lung effusion, and ascites fluid of mammal,  
XX where the neoplasm is derived from endodermal tissue and is selected from  
XX colon cancer, breast cancer, pancreatic cancer, liver cancer, cancer of  
XX bile ducts, and cancer of the CNS. HAAH antibodies, linked to a cytotoxic  
XX agent, are useful for killing tumour cells

SQ Sequence 20 BP; 2 A; 7 C; 7 G; 4 T; 0 U; 0 Other;  
Query Match 0.9%; Score 20; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGACCGTGCATGCCCCAG 20  
DB 20 CGGACCGTGCATGCCCCAG 1  
RESULT 46  
AAF89813/c  
ID AAF89813 standard; DNA; 20 BP.  
XX  
AC AAF89813;  
XX  
DT 23-JUL-2001 (first entry)  
XX  
DE Human aspartyl (asparaginyl) beta-hydroxylase antisense molecule.  
XX  
KW Epidermal growth factor-like domain; EGF-like domain; cancer;  
KW human aspartyl beta-hydroxylase; HAAH; malignant neoplasm; tumour;  
KW antisense oligonucleotide; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200135102-A2.  
XX  
PD 17-MAY-2001.  
XX  
PF 08-NOV-2000; 2000WO-US030738.  
XX  
PR 08-NOV-1999; 99US-00436184.  
XX  
PA (RHOD-) RHODE ISLAND HOSPITAL LIFESPAN PARTNER.  
XX  
PI Wands JR, De La Monte SM, Ince N, Carlson RI;  
XX WPI; 2001-329171/34.  
XX  
PT Diagnosing malignant neoplasm in a mammal, involves contacting mammalian  
PT sample with antibody that binds to human aspartyl beta-hydroxylase  
PT polypeptide to form antigen-antibody complex and detecting the complex.  
XX  
PS Claim 14; Page 57; 76pp; English.  
XX  
CC The present sequence represents an antisense oligonucleotide which  
CC inhibits human aspartyl (asparaginyl) beta-hydroxylase (HAAH) gene.  
CC Epidermal growth factor (EGF)-like domains of polypeptides are  
CC hydroxylated by HAAH enzymes. HAAH is used in the method of the  
CC invention. The specification describes a method for diagnosing a  
CC malignant neoplasm in a mammal. The method comprises contacting a body  
CC fluid with an antibody which binds to HAAH polypeptide under complex  
CC forming conditions, and detecting the antigen-antibody complex. The  
CC method is useful for diagnosing and prognosing a malignant neoplasm in a  
CC bodily fluid e.g. central nervous system (CNS)-derived body fluid, blood,  
CC serum, urine, saliva, sputum, lung effusion, and ascites fluid of mammal,  
CC where the neoplasm is derived from endodermal tissue and is selected from  
CC colon cancer, breast cancer, pancreatic cancer, liver cancer, cancer of  
CC bile ducts, and cancer of the CNS. HAAH antibodies, linked to a cytotoxic  
CC agent, are useful for killing tumour cells  
SQ Sequence 20 BP; 3 A; 6 C; 4 G; 7 T; 0 U; 0 Other;  
Query Match 0.9%; Score 20; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 11 AATGCCCGCAGCGTAAGAATG 30  
DB 20 AATGCCCGCAGCGTAAGAATG 1  
RESULT 48  
ABS57871/c  
ID ABS57871 standard; DNA; 20 BP.  
XX  
AC ABS57871;  
XX  
DT 06-FEB-2003 (first entry)  
XX  
DE Human HAAH antisense oligonucleotide #3.

SQ Sequence 20 BP; 2 A; 7 C; 7 G; 4 T; 0 U; 0 Other;  
Query Match 0.9%; Score 20; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CGGACCGTGCATGCCCCAG 20  
DB 20 CGGACCGTGCATGCCCCAG 1  
RESULT 46  
AAF89813/c  
ID AAF89813 standard; DNA; 20 BP.  
XX  
AC AAF89813;  
XX  
DT 23-JUL-2001 (first entry)  
XX  
DE Human aspartyl (asparaginyl) beta-hydroxylase antisense molecule.  
XX  
KW Epidermal growth factor-like domain; EGF-like domain; cancer;  
KW human aspartyl beta-hydroxylase; HAAH; malignant neoplasm; tumour;  
KW antisense oligonucleotide; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200135102-A2.  
XX  
PD 17-MAY-2001.  
XX  
PF 08-NOV-2000; 2000WO-US030738.  
XX  
PR 08-NOV-1999; 99US-00436184.  
XX  
PA (RHOD-) RHODE ISLAND HOSPITAL LIFESPAN PARTNER.  
XX  
PI Wands JR, De La Monte SM, Ince N, Carlson RI;  
XX WPI; 2001-329171/34.  
XX  
PT Diagnosing malignant neoplasm in a mammal, involves contacting mammalian  
PT sample with antibody that binds to human aspartyl beta-hydroxylase  
PT polypeptide to form antigen-antibody complex and detecting the complex.  
XX  
PS Claim 15; Page 57; 76pp; English.  
XX  
CC The present sequence represents an antisense oligonucleotide which  
CC inhibits human aspartyl (asparaginyl) beta-hydroxylase (HAAH) gene.  
CC Epidermal growth factor (EGF)-like domains of polypeptides are  
CC hydroxylated by HAAH enzymes. HAAH is used in the method of the  
CC invention. The specification describes a method for diagnosing a  
CC malignant neoplasm in a mammal. The method comprises contacting a body  
CC fluid with an antibody which binds to HAAH polypeptide under complex  
CC forming conditions, and detecting the antigen-antibody complex. The  
CC method is useful for diagnosing and prognosing a malignant neoplasm in a  
CC bodily fluid e.g. central nervous system (CNS)-derived body fluid, blood,  
CC serum, urine, saliva, sputum, lung effusion, and ascites fluid of mammal,  
CC where the neoplasm is derived from endodermal tissue and is selected from  
CC colon cancer, breast cancer, pancreatic cancer, liver cancer, cancer of  
CC bile ducts, and cancer of the CNS. HAAH antibodies, linked to a cytotoxic  
CC agent, are useful for killing tumour cells  
SQ Sequence 20 BP; 3 A; 6 C; 6 G; 5 T; 0 U; 0 Other;  
Query Match 0.9%; Score 20; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 6 CGTGCATGCGCCAGCGTAA 25  
DB 20 CGTGCATGCGCCAGCGTAA 1

XX Human; ss; aspartyl (asparaginyl) beta-hydroxylase; HAAH; antisense;  
 KW cytosolic; immunostimulant; antibody; neoplasm; tumour; FB50; 86A; 5C7;  
 KW 19B; brain tumour; glioma; glioblastoma; astrocytoma; haemangioma;  
 KW pancreatic carcinoma; colon cancer; breast cancer; pancreatic cancer;  
 KW liver cancer; cancer of the bile ducts; primary malignant CNS neoplasm;  
 KW metastatic CNS neoplasm.  
 XX Homo sapiens.  
 OS  
 XX US2002110559-A1.  
 XX 15-AUG-2002.  
 XX 17-MAY-2001; 2001US-00859604.  
 XX 08-NOV-1999; 99US-00436184.  
 XX (WAND/) WANDS J R.  
 PA (DMON/) DE LA MONTE S M.  
 PA (DEUT/) DEUTCH A H.  
 PA (GHAN/) GHANBARI H A.  
 XX Wands JR, De La Monte SM, Deutch AH, Ghanbari HA;  
 PI WPI; 2003-066676/06.  
 DR  
 XX Diagnosing neoplasm and inhibiting tumor growth, by contacting tissue of  
 PT mammal with detectably-labeled antibody which binds to human aspartyl  
 PT (asparaginyl) beta-hydroxylase.  
 XX Example 6; Page 18; 34pp; English.  
 PS  
 XX The invention relates to diagnosing a neoplasm and inhibiting tumour  
 CC growth in a mammal, using an antibody that binds to human aspartyl  
 CC (asparaginyl) beta-hydroxylase (HAAH). Diagnosing a neoplasm comprises  
 CC contacting a tissue with a detectably-labelled antibody where an increase  
 CC in level of antibody binding at tissue site compared to the level of  
 CC binding to normal non-neoplastic tissue indicates the presence of a  
 CC neoplasm at the tissue site. Inhibiting tumour growth in mammal involves  
 CC administering the antibody conjugated to a cytotoxic agent to a mammal.  
 CC Also included are a method of conferring an immune response to a tumour  
 CC cell in a mammal, by administering the antibody, a method of inducing an  
 CC HAAH-specific immune response in a mammal, by administering to the mammal  
 CC an HAAH polypeptide (or a polynucleotide composition encoding the  
 CC polypeptide, or its degenerate variant), a fragment of HAAH comprising an  
 CC extracellular domain and lacking a cytoplasmic domain of HAAH, an  
 CC antibody or its fragment which binds to HAAH (where the antibody is FB50,  
 CC 86A, 5C7 or 19B), a hybridoma cell line chosen from hybridoma FB50,  
 CC HA386A, HA15CA and HA219B, and a fragment of HAAH which lacks enzymatic  
 CC activity or alpha-ketoglutarate binding domain and epidermal growth  
 CC factor (EGF)-like domain. The methods are useful for diagnosing neoplasm  
 CC in a mammal, inhibiting tumour growth in a mammal, conferring an immune  
 CC response to a brain tumour cell (e.g. glioma, glioblastoma, astrocytoma  
 CC or haemangioma) in a mammal, for conferring an immune response to a  
 CC pancreatic carcinoma cell and for inducing a HAAH-specific immune  
 CC response in a mammal. The method is useful for diagnosing malignant  
 CC neoplasms derived from endodermal tissue, e.g. colon cancer, breast  
 CC cancer, pancreatic cancer, liver cancer and cancer of the bile ducts. The  
 CC method is also useful for diagnosing neoplasms of central nervous system  
 CC (CNS) e.g. primary malignant CNS neoplasms of both neuronal and glial  
 CC cell origin and metastatic CNS neoplasms, astrocytoma or haemangioma. The  
 CC present sequence represents a Human HAAH antisense oligonucleotide  
 XX Sequence 20 BP; 2 A; 7 C; 7 G; 4 T; 0 U; 0 Other;  
 SQ  
 Query Match 0.9%; Score 20; DB 8; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 CGGACCGTGCATGGCCAG 20  
 |||||

Db 20 CGGACCGTGCATGGCCAG 1  
 RESULT 49  
 ABS57870/c  
 ID ABS57870 standard; DNA; 20 BP.  
 XX ABS57870;  
 AC  
 XX  
 XX 06-FEB-2003 (first entry)  
 DT  
 XX  
 XX Human HAAH antisense oligonucleotide #2.  
 DE  
 XX  
 XX Human; ss; aspartyl (asparaginyl) beta-hydroxylase; HAAH; antisense;  
 KW cytosolic; immunostimulant; antibody; neoplasm; tumour; FB50; 86A; 5C7;  
 KW 19B; brain tumour; glioma; glioblastoma; astrocytoma; haemangioma;  
 KW pancreatic carcinoma; colon cancer; breast cancer; pancreatic cancer;  
 KW liver cancer; cancer of the bile ducts; primary malignant CNS neoplasm;  
 KW metastatic CNS neoplasm.  
 XX Homo sapiens.  
 OS  
 XX US2002110559-A1.  
 XX 15-AUG-2002.  
 XX 17-MAY-2001; 2001US-00859604.  
 XX 08-NOV-1999; 99US-00436184.  
 XX (WAND/) WANDS J R.  
 PA (DMON/) DE LA MONTE S M.  
 PA (DEUT/) DEUTCH A H.  
 PA (GHAN/) GHANBARI H A.  
 XX Wands JR, De La Monte SM, Deutch AH, Ghanbari HA;  
 PI WPI; 2003-066676/06.  
 DR  
 XX Diagnosing neoplasm and inhibiting tumor growth, by contacting tissue of  
 PT mammal with detectably-labeled antibody which binds to human aspartyl  
 PT (asparaginyl) beta-hydroxylase.  
 XX Example 6; Page 18; 34pp; English.  
 PS  
 XX The invention relates to diagnosing a neoplasm and inhibiting tumour  
 CC growth in a mammal, using an antibody that binds to human aspartyl  
 CC (asparaginyl) beta-hydroxylase (HAAH). Diagnosing a neoplasm comprises  
 CC contacting a tissue with a detectably-labelled antibody where an increase  
 CC in level of antibody binding at tissue site compared to the level of  
 CC binding to normal non-neoplastic tissue indicates the presence of a  
 CC neoplasm at the tissue site. Inhibiting tumour growth in mammal involves  
 CC administering the antibody conjugated to a cytotoxic agent to a mammal.  
 CC Also included are a method of conferring an immune response to a tumour  
 CC cell in a mammal, by administering the antibody, a method of inducing an  
 CC HAAH-specific immune response in a mammal, by administering to the mammal  
 CC an HAAH polypeptide (or a polynucleotide composition encoding the  
 CC polypeptide, or its degenerate variant), a fragment of HAAH comprising an  
 CC extracellular domain and lacking a cytoplasmic domain of HAAH, an  
 CC antibody or its fragment which binds to HAAH (where the antibody is FB50,  
 CC 86A, 5C7 or 19B), a hybridoma cell line chosen from hybridoma FB50,  
 CC HA386A, HA15CA and HA219B, and a fragment of HAAH which lacks enzymatic  
 CC activity or alpha-ketoglutarate binding domain and epidermal growth  
 CC factor (EGF)-like domain. The methods are useful for diagnosing neoplasm  
 CC in a mammal, inhibiting tumour growth in a mammal, conferring an immune  
 CC response to a brain tumour cell (e.g. glioma, glioblastoma, astrocytoma  
 CC or haemangioma) in a mammal, for conferring an immune response to a  
 CC pancreatic carcinoma cell and for inducing a HAAH-specific immune  
 CC response in a mammal. The method is useful for diagnosing malignant  
 CC neoplasms derived from endodermal tissue, e.g. colon cancer, breast  
 CC cancer, pancreatic cancer, liver cancer and cancer of the bile ducts. The  
 CC method is also useful for diagnosing neoplasms of central nervous system  
 CC (CNS) e.g. primary malignant CNS neoplasms of both neuronal and glial  
 CC cell origin and metastatic CNS neoplasms, astrocytoma or haemangioma. The  
 CC present sequence represents a Human HAAH antisense oligonucleotide  
 XX Sequence 20 BP; 2 A; 7 C; 7 G; 4 T; 0 U; 0 Other;  
 SQ

CC cell origin and metastatic CNS neoplasms, and for diagnosing brain  
 CC tumours e.g. glioma, glioblastoma, astrocytoma or haemangioma. The  
 CC present sequence represents a Human HAAH antisense oligonucleotide  
 XX  
 SQ Sequence 20 BP; 3 A; 6 C; 6 G; 5 T; 0 U; 0 Other;  
 Query Match 0.9%; Score 20; DB 8; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 CGTCAATGGCCAGCGTAA 25  
 Db 20 CGTCAATGGCCAGCGTAA 1  
 RESULT 50  
 ABS57869/c  
 ID ABS57869 standard; DNA; 20 BP.  
 XX  
 AC ABS57869;  
 XX  
 DT 06-FEB-2003 (first entry)  
 XX  
 DE Human HAAH antisense oligonucleotide #1.  
 XX  
 KW Human; ss; aspartyl (asparaginyl) beta-hydroxylase; HAAH; antisense;  
 KW cytosolic; immunostimulant; antibody; neoplasm; tumour; F950; 86A; 5C7;  
 KW 19B; brain tumour; glioma; glioblastoma; astrocytoma; haemangioma;  
 KW pancreatic carcinoma; colon cancer; breast cancer; pancreatic cancer;  
 KW liver cancer; cancer of the bile ducts; primary malignant CNS neoplasm;  
 KW metastatic CNS neoplasm.  
 XX  
 KW Homo sapiens.  
 OS  
 XX  
 FN US2002110559-A1.  
 XX  
 PD 15-AUG-2002.  
 XX  
 PF 17-MAY-2001; 2001US-00859604.  
 XX  
 PR 08-NOV-1999; 99US-00436184.  
 XX  
 PA (WAND/) WANDS J R.  
 PA (DMON/) DE LA MONTE S M.  
 PA (DEUT/) DEUTCH A H.  
 PA (GHAN/) GHANBARI H A.  
 XX  
 PI Wands JR, De La Monte SM, Deutch AH, Ghanbari HA;  
 XX  
 DR WPI; 2003-066676/06.  
 XX  
 PT Diagnosing neoplasm and inhibiting tumor growth, by contacting tissue of  
 PT mammal with detectably-labeled antibody which binds to human aspartyl  
 PT (asparaginyl) beta-hydroxylase.  
 XX  
 PS Example 6; Page 18; 34pp; English.  
 XX  
 CC The invention relates to diagnosing a neoplasm and inhibiting tumour  
 CC growth in a mammal, using an antibody that binds to human aspartyl  
 CC (asparaginyl) beta-hydroxylase (HAAH). Diagnosing a neoplasm comprises  
 CC contacting a tissue with a detectably-labelled antibody where an increase  
 CC in level of antibody binding at tissue site compared to the level of  
 CC binding to normal non-neoplastic tissue indicates the presence of a  
 CC neoplasm at the tissue site. Inhibiting tumour growth in mammal involves  
 CC administering the antibody conjugated to a cytotoxic agent to a mammal.  
 CC Also included are a method of conferring an immune response to a tumour  
 CC cell in a mammal, by administering the antibody, a method of inducing an  
 CC HAAH-specific immune response in a mammal, by administering to the mammal  
 CC an HAAH polypeptide (or a polynucleotide composition encoding the  
 CC polypeptide, or its degenerate variant), a fragment of HAAH comprising an  
 CC extracellular domain and lacking a cytoplasmic domain of HAAH, an  
 CC antibody or its fragment which binds to HAAH (where the antibody is FB50,  
 CC 86A, 5C7 or 19B), a hybridoma cell line chosen from hybridoma FB50,

CC HA386A, HA15CA and HA219B, and a fragment of HAAH which lacks enzymatic  
 CC activity or alpha-ketoglutarate binding domain and epidermal growth  
 CC factor (EGF)-like domain. The methods are useful for diagnosing neoplasm  
 CC in a mammal, inhibiting tumour growth in a mammal, conferring an immune  
 CC response to a brain tumour cell (e.g. glioma, glioblastoma, astrocytoma  
 CC or haemangioma) in a mammal, for conferring immune response to a  
 CC pancreatic carcinoma cell and for inducing a HAAH-specific immune  
 CC response in a mammal. The method is useful for diagnosing malignant  
 CC neoplasms derived from endodermal tissue, e.g. colon cancer, breast  
 CC cancer, pancreatic cancer, liver cancer and cancer of the bile ducts. The  
 CC method is also useful for diagnosing neoplasms of both neuronal and glial  
 CC (CNS) e.g. primary malignant CNS neoplasms, and for diagnosing brain  
 CC cell origin and metastatic CNS neoplasms, and for diagnosing brain  
 CC tumours e.g. glioma, glioblastoma, astrocytoma or haemangioma. The  
 CC present sequence represents a Human HAAH antisense oligonucleotide  
 XX  
 SQ Sequence 20 BP; 3 A; 6 C; 4 G; 7 T; 0 U; 0 Other;  
 Query Match 0.9%; Score 20; DB 8; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 11 AATGGCCCGCGTAAGAATG 30  
 Db 20 AATGGCCCGCGTAAGAATG 1  
 RESULT 51  
 ADR37507  
 ID ADR37507 standard; RNA; 21 BP.  
 XX  
 AC ADR97507;  
 XX  
 DT 02-DEC-2004 (first entry)  
 XX  
 DE Antisense PPIA siRNA oligomer to knockout apoptosis related gene.  
 XX  
 KW ss; DNA/RNA hybrid; apoptosis; cancer; inflammation; autoimmune;  
 KW neurodegenerative disorder; cytostatic; antiinflammatory; PPIA.  
 KW immunosuppressive; neuroprotective; gene therapy; siRNA; PPIA.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT misc\_feature 20..21  
 FT /note= "Oligo dT bases added to the 3' end to stabilise  
 FT binding"  
 XX  
 FN WO2004078783-A2.  
 XX  
 PD 16-SEP-2004.  
 XX  
 PF 05-MAR-2004; 2004WO-GB000957.  
 XX  
 PR 07-MAR-2003; 2003GB-00005267.  
 XX  
 PA (EIRX-) EIRX THERAPEUTICS LTD.  
 XX  
 PI Murphy FJ, Sheehan DE, Keating KE, Hayes I, Seera L;  
 XX  
 DR WPI; 2004-662402/64.  
 XX  
 PT Identifying an agent that modulates the function of an apoptosis-  
 PT associated polypeptide, useful for diagnosing or treating e.g. cancer,  
 PT comprises comparing the binding of the polypeptide to the candidate agent  
 PT and to a control agent.  
 XX  
 PS Example 5; Page 293; 304pp; English.  
 CC This invention relates to novel agents that modulates the function of  
 CC human apoptosis-associated proteins specified within the specification.

Specifically, it refers to a method for the identification of target genes whose expression is correlated with an early stage in the regulation of apoptosis. The present invention describes a method of contacting either candidate agents or control agents to the target genes and assessing the difference of binding and inhibitory activity, where the candidate agent is selected from a low molecular weight organic molecule, an antibody or fragment thereof, an antisense oligonucleotide, a small inhibitory dRNA, or a ribozyme. As such, the compositions and methods are useful for diagnosing and treating diseases or conditions associated with abnormal apoptosis in mammalian tissue, such as cancer, inflammation, autoimmune or neurodegenerative disorders. Accordingly, they exhibit cytotoxic, antiinflammatory, immunosuppressive and neuroprotective activities. These may also be used for drug screening purposes and in gene therapy. This oligonucleotide sequence is an siRNA oligo used to knockout expression of a human gene identified to be associated with the regulation of apoptosis, given in an exemplification of the invention.

Sequence 21 BP; 7 A; 3 C; 5 G; 2 T; 4 U; 0 Other;

Query Match 0.9%; Score 20; DB 13; Length 21;  
Best Local Similarity 80.0%; Pred. No. 1-7e+02;  
Matches 16; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 1789 ACTGGAGTATCCGAGAT 1808  
DB 1 ACUGGAGUUAUCCGAGAT 20

RESULT 52

ACH75949/c  
ID ACH75949 standard; DNA; 550 BP.

XX ACH75949;

DT 29-JUL-2004 (first entry)

DE Human genome derived single exon probe #9144.

XX Human; probe; ss: gene expression; single exon probe; microarray;  
KW alternative splicing event; genomic alteration.

XX Homo sapiens.

OS US2003194704-A1.

PN 16-OCT-2003.

PP 03-APR-2002; 2002US-00029386.

PR 03-APR-2002; 2002US-00029386.

XX (PENN/) PENN S G.

PA (RANK/) RANK D R.

PA (HANZ/) HANZEL D K.

XX Penn SG, Rank DR, Hanzel DK;

PI WPI; 2004-119264/12.

XX New human genome-derived single exon nucleic acid probes useful for human gene expression analysis, for identifying or characterizing alternative splicing events, for assessing genomic alterations or as tools for surveying tissues.

PS Claim 15; SEQ ID NO 9144; 80pp; English.

XX The invention relates to a nucleic acid probe for measuring human gene expression, comprising any of the 27,400 fully defined nucleotide sequences in the specification, or their complements or fragments, and encoding at least 8 amino acids of any of the 6888 amino acid sequences fully defined in the specification. The probe is a single exon probe that hybridises under high stringency conditions to a nucleic acid molecule

expressed in human cells or tissues. Also included are a spatially-addressable set of single exon nucleic acid probes for measuring human gene expression (comprising a plurality of single exon nucleic acid probes cited above, where each of the plurality of probes is separately and addressably isolatable or amplifiable from the plurality), a single exon microarray for measuring human gene expression, a method of measuring human gene expression, a vector comprising the single exon probe cited above, an ORF-encoded peptide comprising at least 8 contiguous amino acids of any of the above-mentioned amino acid sequences (optionally with conservative amino acid substitutions), an isolated antibody that binds specifically to a peptide cited above, methods of selling and/or licensing single exon probes or microarrays to a customer desiring to measure gene expression, a method of providing human gene expression data by subexpression, and a computer-readable storage medium which contains a database having a plurality of records (each record including data on the expression of a single exon probe cited above. The probe, methods and apparatus are useful in gene expression analysis. The probes may be used as tools for surveying tissues to detect the presence of expressed messages that contain their specific exon, or in constructing genome-derived single exon microarrays. In addition, the probes are used in identifying and characterising alternative splicing events, in detecting and characterising gross alterations in the genomic locus that includes their exon, in assessing smaller genomic alterations, in priming the synthesis of nucleic acids, or in expressing the ORF-encoded peptide. The present sequence is a human single exon probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030194704

XX Sequence 550 BP; 144 A; 121 C; 147 G; 138 T; 0 U; 0 Other;

Query Match 0.9%; Score 20; DB 12; Length 550;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1256 TGCAGACCTGCTGAAGCTGA 1275

DB 280 TGCAGACCTGCTGAAGCTGA 261

RESULT 53

ACH96138/c

ID ACH96138 standard; DNA; 786 BP.

XX ACH96138;

DT 29-JUL-2004 (first entry)

DE Klebsiella pneumoniae polynucleotide seqid 1933.

XX Recombinant expression vector; transcription regulatory element;

KW Klebsiella pneumoniae protein; antibacterial; Vaccine; gene; ds.

XX Klebsiella pneumoniae.

XX US6610836-B1.

XX 26-AUG-2003.

XX 27-JAN-2000; 2000US-00489039.

XX 29-JAN-1999; 99US-0117747P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Breton GL, Osborne M;

XX WPI; 2003-895346/82.

XX P-PSDB; ABO62587.

XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for preparing a vaccine composition against Klebsiella pneumoniae.



```
XX PS Disclosure; SEQ ID NO 1933; 932pp; English.
XX CC
XX CC The invention describes a new isolated nucleic acid encoding a Klebsiella
XX CC pneumoniae polypeptide. Also described are: a recombinant expression
XX CC vector comprising the nucleic acid, operably linked to a transcription
XX CC regulatory element; and a cell comprising the recombinant expression
XX CC vector. The nucleic acid is useful for preparing a vaccine composition
XX CC against Klebsiella pneumoniae. This sequence encodes a Klebsiella
XX CC pneumoniae polypeptide of the invention
XX SQ Sequence 786 BP; 129 A; 220 C; 235 G; 202 T; 0 U; 0 Other;
XX
Query Match 0.9%; Score 20; DB 11; Length 786;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 38 CAGCGGCAACAGCAGCAGCA 57
DB 294 CAGCGGCAACAGCAGCAGCA 275
|||||
RESULT 54
AAT62523
ID AAT62523 standard; DNA; 849 BP.
XX AC
XX AAT62523;
XX DT 06-JUN-1997 (first entry)
XX DE IGSP-hPOMC fusion DNA sequence.
XX KW Analgesic; pain; bioartificial organ; pro-opiomelanocotin; POMC;
XX KW beta-endorphin; CNTF; IGSP-hPOMC; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT exon 1..89 /tag= b
FT 5'UTR 1..32 /tag= a
FT misc_feature 43..186 /tag= f
FT /product= "IGSP region"
FT intron 90..168
FT exon 169..849 /tag= c
FT misc_feature 187..806 /tag= d
FT 3'UTR 807..849 /tag= e
FT /tag= e
XX WO9640959-A1.
XX PD 19-DEC-1996.
XX PF 07-JUN-1996; 96WO-US009629.
XX PF 07-JUN-1995; 95US-00481917.
XX PR (CYTO-) CYTOTHERAPEUTICS INC.
XX PI Saydoff J, Wong S;
XX WI; 1997-087062/08.
XX
Stably transformed cells expressing endorphin, enkephalin and
catecholamine - and artificial organs contg. them, useful for control of
pain, esp. implanted in the CNS.
XX
```

```
PS Example; Page 58; 114pp; English.
XX CC
XX CC A DNA sequence (AAT62523) comprising a fusion between IGSP and human pro-
XX CC opiomelanocortin (POMC) DNA can be utilised in the construction of
XX CC vectors for the expression of endorphin cpds. Cleavage enzymes present in
XX CC host cells can generate a suite of endorphins from POMC, some or all of
XX CC which have analgesic properties. A DNA sequence (see also AAT62524) in
XX CC which the ACTH coding region of the POMC gene is deleted is preferred.
XX CC Host cells transfected with vectors (or a single polycistronic vector)
XX CC encoding analgesic cpds. can be encapsulated to form bioartificial organs
XX CC that are implanted into a patient for control of pain
XX SQ Sequence 849 BP; 176 A; 266 C; 286 G; 121 T; 0 U; 0 Other;
XX
Query Match 0.9%; Score 20; DB 2; Length 849;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 43 GCAACAGCAGCAGCAGCGGC 62
DB 271 GCAACAGCAGCAGCAGCGGC 290
|||||
RESULT 55
ABA19178/c
ID ABA19178 standard; DNA; 900 BP.
XX AC
XX ABA19178;
XX DT 23-JAN-2002 (first entry)
XX DE Human nervous system related polynucleotide SEQ ID NO 11509.
XX KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
XX KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
XX KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
XX KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
XX KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
XX KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
XX KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX OS Homo sapiens.
XX WO200159063-A2.
XX PF 16-AUG-2001.
XX PD 17-JAN-2001; 2001WO-US001334.
XX PF 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 24-FEB-2000; 2000US-0184664P.
XX PR 02-MAR-2000; 2000US-0186350P.
XX PR 16-MAR-2000; 2000US-0189874P.
XX PR 17-MAR-2000; 2000US-019076P.
XX PR 18-APR-2000; 2000US-0198123P.
XX PR 19-MAY-2000; 2000US-0205555P.
XX PR 07-JUN-2000; 2000US-0209467P.
XX PR 28-JUN-2000; 2000US-0214886P.
XX PR 30-JUN-2000; 2000US-0215135P.
XX PR 07-JUL-2000; 2000US-0216647P.
XX PR 07-JUL-2000; 2000US-021680P.
XX PR 11-JUL-2000; 2000US-0217487P.
XX PR 11-JUL-2000; 2000US-0217496P.
XX PR 14-JUL-2000; 2000US-0218290P.
XX PR 26-JUL-2000; 2000US-0220963P.
XX PR 26-JUL-2000; 2000US-0220964P.
XX PR 14-AUG-2000; 2000US-0224518P.
XX PR 14-AUG-2000; 2000US-0224519P.
XX PR 14-AUG-2000; 2000US-0225213P.
XX PR 14-AUG-2000; 2000US-0225214P.
XX PR 14-AUG-2000; 2000US-0225266P.
XX PR 14-AUG-2000; 2000US-0225267P.
```



Best Local Similarity 100.0%; Pred. No. 1.6e+02; Indels 0; Gaps 0; Mismatches 0; Conservative 0;

Qy 43 GCAACAGCAGCAGCAGCGGC 62  
Db 764 GCAACAGCAGCAGCAGCGGC 745

RESULT 56  
AAC58389  
ID AAC58389 standard; CDNA; 908 BP.

XX AC AAC58389;

XX DT 29-JAN-2001 (first entry)

XX DE Human PRO23 nucleotide sequence SEQ ID NO:64.

XX KW Human; tumour; diagnosis; neoplastic disease; neoplastic cell growth;  
XX KW proliferation; tumorigenesis; identification; cancer; cytostatic;  
XX KW neurotropic; neuroprotective; antiinflammatory; immunosuppressive;  
XX KW immunostimulant; antiangiogenic; leukaemia; lymphoid malignancy;  
XX KW neuronal disorder; glioma; astrocytic disorder; angiogenic;  
XX KW hypothalamic disorder; glandular disorder; macrophagal disorder;  
XX KW epithelial disorder; stromal disorder; blastocoele disorder;  
XX KW inflammatory disorder; immunologic disorder; ss.

XX OS Homo sapiens.

XX PN WO200053755-A2.

XX PD 14-SEP-2000.

XX PF 06-JAN-2000; 2000WO-US000376.

XX PR 08-MAR-1999; 99WO-US005028.

XX PR 02-JUN-1999; 99WO-US012252.

XX PR 23-JUN-1999; 98US-0141037P.

XX PR 07-JUL-1999; 98US-0143048P.

XX PR 26-JUL-1999; 99US-0145698P.

XX PR 30-NOV-1999; 99WO-US028313.

XX PR 20-DEC-1999; 99WO-US030911.

XX PR 05-JAN-2000; 2000WO-US000219.

XX PA (GETH ) GENENTECH INC.

XX PI Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hillian KJ, Roy MA;  
XX PI Watanabe CK, Wood WI;

XX DR WPI; 2000-572270/53.

XX DR P-PSDB; AAB24079.

XX PT Thirty PRO polynucleotides encoding PRO polypeptides, useful in the  
XX PT treatment, diagnosis and prevention of cancer.

XX PS Claim 50; Fig 45; 286pp; English.

XX CC The present invention describes an isolated antibody that binds to one of  
XX CC the human PRO proteins designated PRO212, PRO290, PRO341, PRO535, PRO619,  
XX CC PRO171, PRO809, PRO830, PRO848, PRO943, PRO1005, PRO1009, PRO1025,  
XX CC PRO1030, PRO1097, PRO1107, PRO1111, PRO1153, PRO1182, PRO1184, PRO1187,  
XX CC PRO1281, PRO1293, PRO339, PRO834, PRO1317, PRO1710, PRO2094, PRO2145 OR  
XX CC PRO2198. PRO antagonists can be used to inhibit tumour cell growth. The  
XX CC PRO polypeptides and nucleotides are useful in the treatment, diagnosis  
XX CC and prevention of cancer. The antibodies and other anti-tumour compounds  
XX CC may be used to treat various conditions, including those characterised by  
XX CC overexpression and/or activation of the amplified PRO genes. Exemplary  
XX CC conditions or disorders to be treated with such antibodies and other  
XX CC compounds include benign or malignant tumours (e.g., renal, liver,  
XX CC kidney, bladder, breast, gastric, ovarian, colorectal, prostate,  
XX CC pancreatic, lung, vulva, thyroid, hepatic carcinomas, sarcomas,  
XX CC glioblastomas, and various head and neck tumours), leukaemias and  
XX CC lymphoid malignancies, other disorders such as neuronal, glial,

CC astrocytal, hypothalamic and other glandular, macrophagal, epithelial,  
CC stromal and blastocoele disorders, and inflammatory, angiogenic and  
CC immunologic disorders. AAC58242 to AAC58366 represent PCR primers and  
CC hybridisation probes used in the isolation of the human PRO sequences.  
CC AAC58367 to AAC58396 and AAB24057 to AAB24089 represent human PRO  
CC polynucleotide and protein sequences given in the exemplification of the  
XX CC present invention

SQ Sequence 908 BP; 173 A; 308 C; 310 G; 117 T; 0 U; 0 Other;

Query Match 0.9%; Score 20; DB 3; Length 908;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCAGCGGC 62

Db 254 GCAACAGCAGCAGCAGCGGC 273

RESULT 57

AAL49917

ID AAL49917 standard; DNA; 936 BP.

XX AC AAL49917;

XX DT 05-DEC-2002 (first entry)

XX DE Human pro-opiomelanocortin coding sequence fragment.

XX KW Human; pro-opiomelanocortin; POMC; corticosteroid; pain; analgesic;  
XX KW chronic inflammation; neuroendocrine; CRF; beta-endorphin; gene therapy;  
XX KW corticotrophin-releasing factor; antiinflammatory; gene; ds.

XX OS Homo sapiens.

XX PN WO200267996-A2.

XX PD 06-SEP-2002.

XX PF 19-FEB-2002; 2002WO-DE000583.

XX PR 24-FEB-2001; 2001DE-01009092.

XX PA (MOLO-) MOLOGEN FORSCH ENTWICKLUNGS & VERTRIEBS.

XX PI Wittig B, Stein C, Schaefer M, Schroff M, Junghans C;  
XX PI Koenig Merediz SA;

XX DR WPI; 2002-682796/73.

XX PT Agent for reducing pain, useful particularly for chronic inflammatory  
XX PT pain, comprises a nucleic acid construct for local expression of  
XX PT neuropeptides.

XX PS Claim 22; Page 33; 34pp; German.

XX CC The present invention relates to an agent for reducing or suppressing  
XX CC pain in higher animals, especially humans, which comprises (free from  
XX CC cells or cell lysate) an expression construct for local expression of a  
XX CC neuroendocrine peptide. The agent is especially useful for treating pain  
XX CC associated with chronic inflammatory disorders. The neuroendocrine may be  
XX CC pro-opiomelanocortin (POMC) or corticotrophin-releasing factor (CRF). The  
XX CC present sequence is a fragment of the human POMC coding sequence

SQ Sequence 936 BP; 190 A; 297 C; 328 G; 121 T; 0 U; 0 Other;

Query Match 0.9%; Score 20; DB 6; Length 936;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCAGCGGC 62

Db 401 GCAACAGCAGCAGCAGCGGC 420

```
RESULT 58
ABL27131
ID ABL27131 standard; DNA; 984 BP.
XX
AC ABL27131;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 32866.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
FN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
PI WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 32866; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 984 BP; 332 A; 268 C; 225 G; 159 T; 0 U; 0 Other;
Query Match 0.9%; Score 20; DB 4; Length 984;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 38 CAGCGGCAACAGCAGCAGCA 57
DB 825 CAGCGGCAACAGCAGCAGCA 844
RESULT 59
ACN42014
ID ACN42014 standard; cDNA; 1038 BP.
XX
AC ACN42014;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human diagnostic and therapeutic polynucleotide SEQ ID NO:889.
XX
KW ss; gene; gene therapy; human diagnostic and therapeutic polynucleotide;
KW dithp.
XX
OS Homo sapiens.
XX
FN WO2004023973-A2.
XX
PD 25-MAR-2004.
XX
PF 12-SEP-2003; 2003WO-US028227.
XX
PR 12-SEP-2002; 2002US-0410259P.
XX
PR 12-SEP-2002; 2002US-0410260P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;
PI Harthorne TA, Suchorski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Deleage AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstein EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX
WPI; 2004-329368/30.
XX
P-PSDB; ABM83362.
XX
PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
PS Claim 1; Page; 190pp; English.
XX
CC The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders, endocrine
CC autoimmune/inflammatory disorders, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp polynucleotide of
CC the invention. Note: The sequence data for this patent is not represented
CC in the printed specification, but was obtained in electronic format
CC directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ Sequence 1038 BP; 208 A; 345 C; 355 G; 130 T; 0 U; 0 Other;
Query Match 0.9%; Score 20; DB 13; Length 1038;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 43 GCAACAGCAGCAGCAGCGGC 62
DB 414 GCAACAGCAGCAGCAGCGGC 433
RESULT 60
ACN42013
ID ACN42013 standard; cDNA; 1039 BP.
XX
AC ACN42013;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human diagnostic and therapeutic polynucleotide SEQ ID NO:888.
XX
KW ss; gene; gene therapy; human diagnostic and therapeutic polynucleotide;
KW dithp.
XX
```

OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
XX 27-SEP-2001.  
PD  
XX 23-MAR-2001; 2001WO-US009231.  
XX  
XX 23-MAR-2000; 2000US-0191637P.  
XX  
PR 11-JUL-2000; 2000US-00614150.  
XX  
XX (PEKE ) PE CORP NY.  
PA  
XX Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
XX  
XX Claim 1; SEQ ID NO 1798; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 1070 BP; 291 A; 321 C; 277 G; 181 T; 0 U; 0 Other;  
SQ  
Query Match 0.9%; Score 20; DB 4; Length 1070;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 37 GCAGCGGCAACAGCAGCAGC 56  
DB 700 GCAGCGGCAACAGCAGCAGC 681  
RESULT 62  
ACN42012  
ID ACN42012 standard; cDNA; 1088 BP.  
XX  
XX ACN42012;  
XX  
XX 18-NOV-2004 (first entry)  
XX Human diagnostic and therapeutic polynucleotide SEQ ID NO:887.  
XX  
XX ss; gene; gene therapy; human diagnostic and therapeutic polynucleotide;  
XX dithp.  
XX Homo sapiens.  
XX  
XX WO2004023973-A2.  
XX  
XX 25-MAR-2004.  
PD  
XX 12-SEP-2003; 2003WO-US028227.  
XX  
XX 12-SEP-2002; 2002US-0410259P.  
XX  
PR 12-SEP-2002; 2002US-0410260P.  
XX  
XX (INCY-) INCYTE CORP.  
XX  
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;  
XX Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;  
XX Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;  
XX Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;  
XX Petalita CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;  
XX Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;  
XX Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;  
XX Patry S, Shi X, Suarez CJ;  
XX  
XX WPI; 2004-329368/30.  
XX P-PSDB; ABM83361.  
XX  
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful  
PT in diagnosing a condition, disease or disorder associated with human  
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or  
PT in gene mapping.  
XX  
XX Claim 1; Page; 190pp; English.  
XX  
XX The invention relates to novel diagnostic and therapeutic polynucleotides  
CC selected from one of the 2722 sequences defined in the specification. A  
CC polynucleotide of the invention may have a use in gene therapy. The human  
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be  
CC used to diagnose a particular condition, disease or disorder associated  
CC with human molecules, e.g. cell proliferative disorders,  
CC autoimmune/inflammatory disorder, developmental disorder, endocrine  
CC disorder, neurological disorders, gastrointestinal disorders, or  
CC infections caused by virus, bacteria, fungi or parasite. The dithp  
CC molecules may also be used in genetic mapping, in identifying individuals  
CC from minute biological samples, in detecting single nucleotide  
CC polymorphisms, as molecular weight markers, and for somatic or germline  
CC gene therapy. The present sequence represents a dithp polynucleotide of  
CC the invention. Note: The sequence data for this patent is not represented  
CC in the printed specification, but was obtained in electronic format  
CC directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm  
XX  
XX Sequence 1039 BP; 218 A; 337 C; 347 G; 137 T; 0 U; 0 Other;  
SQ  
Query Match 0.9%; Score 20; DB 13; Length 1039;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 43 GCAACAGCAGCAGCAGCAGC 62  
DB 414 GCAACAGCAGCAGCAGCAGC 433  
RESULT 61  
ABL16775/c  
ID ABL16775 standard; DNA; 1070 BP.  
XX  
XX ABL16775;  
XX  
XX 26-MAR-2002 (first entry)  
XX  
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 1798.  
XX  
XX Drosophila; developmental biology; cell signalling; insecticide;  
XX pharmaceutical; gene; ds.  
XX  
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;  
XX Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;

OS Homo sapiens.  
XX  
PN WO2004023973-A2.  
XX  
XX 25-MAR-2004.  
PD  
XX 12-SEP-2003; 2003WO-US028227.  
XX  
XX 12-SEP-2002; 2002US-0410259P.  
XX  
PR 12-SEP-2002; 2002US-0410260P.  
XX  
XX (INCY-) INCYTE CORP.  
XX  
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;  
XX Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;  
XX Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;  
XX Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;  
XX Petalita CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;  
XX Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;  
XX Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;  
XX Patry S, Shi X, Suarez CJ;  
XX  
XX WPI; 2004-329368/30.  
XX P-PSDB; ABM83361.  
XX  
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful  
PT in diagnosing a condition, disease or disorder associated with human  
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or  
PT in gene mapping.  
XX  
XX Claim 1; Page; 190pp; English.  
XX  
XX The invention relates to novel diagnostic and therapeutic polynucleotides  
CC selected from one of the 2722 sequences defined in the specification. A  
CC polynucleotide of the invention may have a use in gene therapy. The human  
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be  
CC used to diagnose a particular condition, disease or disorder associated  
CC with human molecules, e.g. cell proliferative disorders,  
CC autoimmune/inflammatory disorder, developmental disorder, endocrine  
CC disorder, neurological disorders, gastrointestinal disorders, or  
CC infections caused by virus, bacteria, fungi or parasite. The dithp  
CC molecules may also be used in genetic mapping, in identifying individuals  
CC from minute biological samples, in detecting single nucleotide  
CC polymorphisms, as molecular weight markers, and for somatic or germline  
CC gene therapy. The present sequence represents a dithp polynucleotide of  
CC the invention. Note: The sequence data for this patent is not represented  
CC in the printed specification, but was obtained in electronic format  
CC directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm  
XX  
XX Sequence 1039 BP; 218 A; 337 C; 347 G; 137 T; 0 U; 0 Other;  
SQ  
Query Match 0.9%; Score 20; DB 13; Length 1039;  
Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 43 GCAACAGCAGCAGCAGCAGC 62  
DB 414 GCAACAGCAGCAGCAGCAGC 433  
RESULT 61  
ABL16775/c  
ID ABL16775 standard; DNA; 1070 BP.  
XX  
XX ABL16775;  
XX  
XX 26-MAR-2002 (first entry)  
XX  
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 1798.  
XX  
XX Drosophila; developmental biology; cell signalling; insecticide;  
XX pharmaceutical; gene; ds.  
XX  
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;  
XX Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;



PR 19-MAR-1998; 98US-0078566P.  
 PR 19-MAR-1998; 98US-0078573P.  
 PR 19-MAR-1998; 98US-0078574P.  
 PR 19-MAR-1998; 98US-0078576P.  
 PR 19-MAR-1998; 98US-0078577P.  
 PR 19-MAR-1998; 98US-0078578P.  
 PR 19-MAR-1998; 98US-0078579P.  
 PR 19-MAR-1998; 98US-0078581P.  
 PR 01-APR-1998; 98US-0080312P.  
 PR 01-APR-1998; 98US-0080313P.  
 PR 01-APR-1998; 98US-0080314P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Ruben SM, Ni J, Rosen CA, Yu G, Young PE, Feng P, Soppet DR;  
 PI Wei Y, Endress GA, Duan RD, Kyaw H, Ebner R, Lafleur DW, Olsen HS;  
 PI Shi Y, Moore PA;  
 XX WPI; 1999-562050/47.  
 DR P-PSDB; AAY41377.  
 XX  
 XX New isolated human genes, useful for diagnosis and treatment of e.g.  
 PT cancers, neurological disorders, immune diseases, inflammation or blood  
 PT disorders.  
 XX  
 XX Claim 1; Page 340-341; 484pp; English.  
 XX  
 XX This sequence represents a nucleic acid molecule which encodes a secreted  
 CC human protein. The gene number, and the clone it is derived from, are  
 CC detailed in the descriptor line. The gene can be used to generate fusion  
 CC proteins by linking to the gene to a human immunoglobulin Fc portion  
 CC (e.g. AAZ24802) for increasing the stability of the fused protein as  
 CC compared to the human protein only. The invention relates to 95 novel  
 CC genes and their fragments (nucleic acid sequences: AAZ24811-224907; amino  
 CC acid sequences AAY41308-Y41404) which are useful for preventing, treating  
 CC or ameliorating medical conditions e.g. by protein or gene therapy. Also,  
 CC pathological conditions can be diagnosed by determining the amount of the  
 CC new polypeptides in a sample or by determining the presence of mutations  
 CC in the new polynucleotides. Specific uses are described for each of the  
 CC 95 polynucleotides, based on which tissues they are most highly expressed  
 CC in (see AAZ24811 for described uses)  
 XX  
 XX Sequence 1109 BP; 103 A; 454 C; 387 G; 165 T; 0 U; 0 Other;  
 SQ  
 Query Match 0.9%; Score 20; DB 2; Length 1109;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 43 GCAACAGCAGCAGCAGCGGC 62  
 Db 152 GCAACAGCAGCAGCAGCGGC 133  
 RESULT 65  
 AAA13712  
 ID AAA13712 standard; DNA; 1230 BP.  
 XX  
 XX AAA13712;  
 AC  
 XX 25-JUL-2000 (first entry)  
 DT  
 XX Human proopiomelanocortin gene exon 3 nucleotide sequence.  
 DE  
 XX Human; proopiomelanocortin; POMC; beta-endorphin; opioid; addiction;  
 KW amphetamine; nervous system; ds.  
 KW  
 XX Homo sapiens.  
 OS  
 XX NO9804992-A.  
 PN  
 XX 28-APR-2000.  
 FD  
 XX 27-OCT-1998; 98NO-00004992.  
 PF

XX 27-OCT-1998; 98NO-00004992.  
 PR (GENO-) GENOVA AS.  
 XX  
 XX Fagerlund TH, Alestrom P, Berg K;  
 PI  
 XX WPI; 2000-373908/32.  
 DR  
 XX Novel isolated and purified nucleic acid molecule encoding beta-  
 PT endorphin, useful for treating a patient for psychological dependence on,  
 PT and abuse of, substances which have a stimulating and euphoric effect.  
 XX  
 XX Disclosure; Page 3-4; 8pp; Norwegian.  
 XX  
 XX The present invention describes recombinant plasmids for peripheral and  
 CC central nervous system analysing abuse of euphoric and stimulation  
 CC agents. The present sequence represents the nucleotide sequence of exon 3  
 CC from the human proopiomelanocortin (POMC) gene, which is used in the  
 CC exemplification of the present invention  
 XX  
 XX Sequence 1230 BP; 265 A; 379 C; 411 G; 175 T; 0 U; 0 Other;  
 SQ  
 Query Match 0.9%; Score 20; DB 3; Length 1230;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 43 GCAACAGCAGCAGCAGCGGC 62  
 Db 474 GCAACAGCAGCAGCAGCGGC 493  
 RESULT 66  
 ADE28660/C  
 ID ADE28660 standard; cDNA; 1245 BP.  
 XX  
 XX ADE28660;  
 AC  
 XX 29-JAN-2004 (first entry)  
 DT  
 XX Human NOV11a cDNA - SEQ ID 37.  
 DE  
 XX  
 XX NOVX; antidiabetic; anorectic; cardiatic; hypotensive;  
 KW antiarteriosclerotic; virucide; antibacterial; fungicide; protozoacide;  
 KW norectic; neuroprotective; antiparkinsonian; anticonvulsant;  
 KW osteopathic; antiarthritic; antiinflammatory; dermatological;  
 KW anorexia; cancer; cardiovascular; hypertension; atherosclerosis;  
 KW neurodegenerative; Alzheimer's disease; Parkinson's; epilepsy; immune;  
 KW osteoarthritis; haemopoietic; inflammatory skin; asthma; dyslipidaemia;  
 KW neurogenesis; cell differentiation; proliferation; haemopoiesis;  
 KW wound healing; angiogenesis; gene therapy; chromosome mapping;  
 KW tissue typing; human; ss; gene; NOV.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO2003040330-A2.  
 PN  
 XX 15-MAY-2003.  
 PD  
 XX 05-NOV-2002; 2002WO-US035536.  
 PF  
 XX 05-NOV-2001; 2001US-0338626P.  
 PR 05-DEC-2001; 2001US-0336600P.  
 PR 07-DEC-2001; 2001US-0338285P.  
 PR 12-DEC-2001; 2001US-0341346P.  
 PR 17-DEC-2001; 2001US-0341477P.  
 PR 17-DEC-2001; 2001US-0341540P.  
 PR 20-DEC-2001; 2001US-0342592P.  
 PR 27-DEC-2001; 2001US-0344297P.  
 PR 31-DEC-2001; 2001US-0344903P.  
 PR 17-APR-2002; 2002US-0373288P.  
 PR 15-MAY-2002; 2002US-0380981P.



PR 17-MAY-2002; 2002US-0381495P.  
 PR 28-MAY-2002; 2002US-0383534P.  
 PR 28-MAY-2002; 2002US-0383744P.  
 PR 29-MAY-2002; 2002US-0383829P.  
 PR 29-MAY-2002; 2002US-0384024P.  
 PR 07-AUG-2002; 2002US-0401788P.  
 PR 26-AUG-2002; 2002US-0406353P.  
 PR 31-OCT-2002; 2002US-00287971.  
 XX  
 XX  
 XX (CURA-) CURAGEN CORP.  
 XX  
 XX Alsobrook JP, Alvarez E, Anderson DW, Baron M, Boldog FL, Eisen A;  
 PI Burgess CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A;  
 PI Ellerman K, Ettenberg S, Gangolli EA, Gerlach VL, Gorman L;  
 PI Grosse WM, Guo X, Hackett C, Ji W, Kekuda R, Khrantsov NV;  
 PI Lepley DM, Li L, MacDougall JR, Malyankar UM, Mazur A, McQueeney K;  
 PI Mezes PS, Miller CE, Millet I, Mishra VS, Padigar M, Patturajan M;  
 PI Pena CE, Peyman JA, Rastelli L, Rieger DK, Shenoy SG, Shinkets RA;  
 PI Smithson G, Starling G, Szytek KA, Stone DJ, Tchernev VT, Twomlow N;  
 PI Vernet CAM, Zerhusen BD, Zhong M;  
 XX  
 DR WPI: 2003-441555/41.  
 DR P-PSDB; ADE28661.  
 XX  
 XX New isolated NOVX polypeptides and polynucleotides, useful for  
 PT preventing, diagnosing or treating NOVX-associated disorders, e.g.  
 PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,  
 PT asthma, or infections.  
 XX  
 XX Claim 20; SEQ ID NO 37; 447pp; English.  
 XX  
 XX The invention relates to a novel isolated NOVX polypeptide. The  
 CC polypeptide of the invention demonstrates, antidiabetic, anorectic,  
 CC radiant, hypotensive, antiarteriosclerotic, virucide, antibacterial,  
 CC fungicide, protoscidic, nootropic, neuroprotective, antiparkinsonian,  
 CC anticonvulsant, osteopathic, antiarthritic, antiinflammatory,  
 CC dermatological, antiaesthetic and antilipemic activities. The  
 CC polypeptides, nucleic acid molecules and antibodies may be useful for  
 CC treating or diagnosing diseases including metabolic disorders such as  
 CC diabetes and obesity, infectious diseases, anorexia, cancer,  
 CC cardiovascular diseases including hypertension and atherosclerosis,  
 CC neurodegenerative disorders such as Alzheimer's disease, Parkinson's  
 CC disease and epilepsy, immune disorders e.g. osteoarthritis, haemopoietic  
 CC disorders, inflammatory skin disorders, asthma and dyslipidaemia.  
 CC Furthermore, the nucleic acids and polypeptides may also be used to  
 CC identify molecules that modulate or inhibit neurogenesis, cell  
 CC differentiation and proliferation, haemopoiesis, wound healing and  
 CC angiogenesis, as well as in gene therapy. Finally, the nucleic acids may  
 CC be used as hybridisation probes, in chromosome mapping, tissue typing,  
 CC preventive medicine and pharmacogenomics. The current sequence is that  
 CC of the human NOVX cDNA of the invention.  
 XX  
 SQ Sequence 1245 BP; 170 A; 486 C; 407 G; 182 T; 0 U; 0 Other;  
 Query Match 0.9%; Score 20; DB 10; Length 1245;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 43 GCACACGACGACGACGCGC 62  
 DB 138 GCACACGACGACGACGCGC 119  
 RESULT 67  
 ID ADM93405/c  
 XX ID ADM93405 standard; cDNA; 1245 BP.  
 AC ADM93405;  
 XX  
 XX 01-JUL-2004 (first entry)  
 XX Human NOVX polynucleotide #19.  
 DE  
 XX

KW Human; NOVX; gene; ss; congenital heart defect; cardiomyopathy;  
 KW atherosclerosis; hypertension; pulmonary stenosis; scleroderma;  
 KW adenocarcinoma; haemophilia; graft-versus-host disease; cancer;  
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
 KW multiple sclerosis; diabetes; obesity; bronchial asthma;  
 KW acquired immunodeficiency syndrome; AIDS; Crohn's disease;  
 KW infectious disease; anorexia; immune disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 XX US2004067882-A1.  
 XX  
 XX 08-APR-2004.  
 XX  
 XX 05-NOV-2002; 2002US-00287971.  
 XX  
 XX 22-OCT-2001; 2001US-00035568.  
 XX 05-NOV-2001; 2001US-0386266P.  
 XX 06-NOV-2001; 2001US-033072P.  
 XX 09-NOV-2001; 2001US-0345398P.  
 XX 09-NOV-2001; 2001US-0348283P.  
 XX 15-NOV-2001; 2001US-0335610P.  
 XX 21-NOV-2001; 2001US-0332152P.  
 XX 28-NOV-2001; 2001US-0333912P.  
 XX 29-NOV-2001; 2001US-00997425.  
 XX 29-NOV-2001; 2001US-0334300P.  
 XX 04-DEC-2001; 2001US-033576P.  
 XX 05-FEB-2002; 2002US-0354807P.  
 XX 15-MAY-2002; 2002US-0380968P.  
 XX 16-MAY-2002; 2002US-0381043P.  
 XX 02-JUL-2002; 2002US-0393148P.  
 XX 02-JUL-2002; 2002US-0393262P.  
 XX 06-AUG-2002; 2002US-0401479P.  
 XX 06-AUG-2002; 2002US-0401626P.  
 XX 07-AUG-2002; 2002US-0401593P.  
 XX 07-AUG-2002; 2002US-0401695P.  
 XX 26-AUG-2002; 2002US-0406181P.  
 XX  
 XX (ALSO/) ALSOBROOK J P.  
 XX (ALVA/) ALVAREZ E.  
 XX (ANDE/) ANDERSON D W.  
 XX (BARO/) BARON M.  
 XX (BOLD/) BOLDOG F L.  
 XX (BURG/) BURGESS C E.  
 XX (CASW/) CASMAN S J.  
 XX (CHAP/) CHAPOVAL A.  
 XX (DHAN/) DHANABAL M.  
 XX (EDIN/) EDINGER S R.  
 XX (EISE/) EISEN A.  
 XX (ELLE/) ELLERMAN K.  
 XX (ETTE/) ETTEMBERG S.  
 XX (GANG/) GANGOLLI E A.  
 XX (GERL/) GERLACH V.  
 XX (GORM/) GORMAN L.  
 XX (GROS/) GROSSE W M.  
 XX (GUOX/) GUO X.  
 XX (HACK/) HACKETT C.  
 XX (JIW/) JI W.  
 XX (KEKU/) KEKUDA R.  
 XX (KHRA/) KHRANTSOV N V.  
 XX (LEPL/) LEPLEY D M.  
 XX (LILL/) LI L.  
 XX (MACD/) MACDOUGALL J R.  
 XX (MALY/) MALYANKAR U M.  
 XX (MAZU/) MAZUR A.  
 XX (MCQU/) MCQUEENEY K.  
 XX (MEZE/) MEZES P S.  
 XX (MILL/) MILLER C E.  
 XX (MILM/) MILLET I.  
 XX (MISH/) MISHRA V.  
 XX (PADI/) PADIGARU M.  
 XX (PATI/) PATTURAJAN M.  
 XX (PENA/) PENNA C E A.

PA (PEYM/) PEYMAN J A.  
PA (RAST/) RASTELLI L.  
PA (RIEG/) RIEGER D K.  
PA (ROTH/) ROTHENBERG M E.  
PA (SHEN/) SHENOY S G.  
PA (SHIM/) SHIMKETS R A.  
PA (SMIT/) SMITHSON G.  
PA (SPAD/) SPADERNA S K.  
PA (STAR/) STARLING G.  
PA (SPYT/) SPYTEK K A.  
PA (STON/) STONE D J.  
PA (TCHE/) TCHERNEV V T.  
PA (TWO/) TWOMLOW N.  
PA (VERN/) VERNET C A M.  
PA (ZERN/) ZERNHUSEN B D.  
PA (VOSS/) VOSS E Z.  
PA (ZHON/) ZHONG M.  
XX  
PI Alsobrook JP, Alvarez E, Anderson DW, Baron M, Boldog FL, Burgeas CE, Casman SJ, Chapoval A, Dhanabal M, Edinger SR, Eisen A, Ellerman K, Etenberg S, Gangoli EA, Gerlach V, Gorman L, Grosse WM, Guo X, Hackett C, Ji W, Kekuda R, Khrantsov NV, Lepley DW, Li L, Macdougall JR, Malyankar UM, Mazur A, McQueeney K, Mezes PS, Miller CE, Millet I, Mishra V, Padigaru M, Patturajan M, Pena CEA, Peyman JA, Rastelli L, Rieger DK, Rothenberg ME, Shenoy SG, Shimkets RA, Smithson G, Spaderna SK, Starling G, Spytok KA, Stone DJ, Tchernev VT, Twomlow N, Vernet CAM, Zernhuseen BD, Voss EZ, Zhong M;  
XX WPI: 2004-355303/33.  
DR P-PSDB; ADM93406.  
XX  
XX Novel isolated NOVX polypeptide useful treating or preventing disorders or syndromes such as Alzheimer's disease, Parkinson's disease, multiple sclerosis, diabetes, obesity, cancer, bronchial asthma, Crohn's disease.  
XX Claim 20; SEQ ID NO 37; 330bp; English.  
XX  
XX The invention relates to human NOVX polypeptides and the polynucleotides encoding them. The NOVX polypeptides and polynucleotides are useful for determining the presence of or predisposition to a disease associated with altered levels of the sequences in a mammalian subject, and for treating or preventing a pathology associated with NOVX. The polypeptides, polynucleotides and antibodies that bind immunospecifically to the polypeptides are useful for treating or preventing disorders or syndromes such as congenital heart defects, cardiomyopathy, atherosclerosis, hypertension, pulmonary stenosis, scleroderma, adenocarcinoma, haemophilia, graft-versus-host disease, cancer, neurodegenerative disorders, Alzheimer's disease, Parkinson's disease, multiple sclerosis, diabetes, obesity, bronchial asthma, acquired immunodeficiency syndrome (AIDS), Crohn's disease, infectious disease, anorexia and immune disorders. This sequence represents a human NOVX polynucleotide of the invention. Note: The sequence data for this patent is also available from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
XX  
SQ Sequence 1245 BP; 170 A; 486 C; 407 G; 182 T; 0 U; 0 Other;  
Query Match 0.9%; Score 20; DB 12; Length 1245;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 43 GCACAGCAGCAGCAGCGGC 62  
Db 138 GCACAGCAGCAGCAGCGGC 119  
RESULT 68  
ACC79031/c  
ID ACC79031 standard; cDNA; 1418 BP.  
XX  
XX ACC79031;  
XX  
DT 30-JUL-2003 (first entry)

XX DE  
XX  
XX KW  
KW Human; secreted protein; SECP; cytostatic; antiarteriosclerotic; anticonvulsant; neurotropic; neuroprotective; cerebroprotective; anti-HIV; anti-allergic; anti-inflammatory; thymimetic; gene therapy; cancer; cell proliferative disorder; atherosclerosis; neurological disorder; epilepsy; Huntington's disease; stroke; immune disorder; allergy; AIDS; inflammatory disorder; developmental disorder; hypothyroidism; Cushing's syndrome; infection; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO2003016506-A2.  
XX  
XX 27-FEB-2003.  
XX  
XX 15-AUG-2002; 2002WO-US027143.  
XX  
XX 17-AUG-2001; 2001US-0313249P.  
PR 24-AUG-2001; 2001US-0314752P.  
PR 07-SEP-2001; 2001US-0317818P.  
PR 07-SEP-2001; 2001US-0317824P.  
PR 21-SEP-2001; 2001US-0324040P.  
PR 24-SEP-2001; 2001US-0324586P.  
PR 02-NOV-2001; 2001US-0343980P.  
PR 28-NOV-2001; 2001US-0334229P.  
PR 13-FEB-2002; 2002US-0357002P.  
PR 06-MAR-2002; 2002US-0362439P.  
PR 19-MAR-2002; 2002US-0366041P.  
PR 30-APR-2002; 2002US-0376988P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Tang TY, Warren BA, Gietzen KJ, Lal PG, Yue H, Honchell CD; Lehr-Mason PM, Burford N, Xu Y, Baughn MR, Duggan BM, Tran UK; Lee EA, Forsythe IJ, Richardson TW, Lee S, Thangavelu K, Yue H; Emerling BM, Walla NK, Azimzai Y, Sanjanwala B, Hafalia AJA; PI Borowsky ML, Nguyen DB, Ison CH, Astronoff A, Ding L, Lee SY; Becha SD, Ramkumar J, Gandhi AR, Jin P, Fu GK, Swarnakar A;  
XX WPI: 2003-278569/27.  
DR P-PSDB; ABR56731.  
XX  
XX New human secreted proteins (SECP), useful for diagnosing, treating and preventing diseases or conditions associated with the aberrant SECP expression e.g. cancer, AIDS, atherosclerosis, Huntington's disease, stroke, infections.  
XX  
XX Claim 5; Page 246; 286pp; English.  
XX  
XX ACC79026 to ACC79105 encode the human secreted proteins (I) given in ABR56726 to ABR56805, designated SECP-1 to SECP-80. SECP sequences can have cytostatic, antiarteriosclerotic, anticonvulsant, anti-inflammatory, neurotropic, neuroprotective, cerebroprotective, anti-HIV, anti-allergic and anticonvulsant activities, and can be used in gene therapy. The SECP proteins and polynucleotides can be used in diagnosing, treating and preventing diseases or conditions associated with the decreased expression or overexpression of SECP, such as cell proliferative (e.g. cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS, allergies) and developmental (e.g. hypothyroidism, Cushing's syndrome) disorders, or infections. They are also useful in assessing the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of SECP. The SECP or its fragments are useful in screening compounds for effectiveness as agonist or antagonist of the polypeptides, or in altering the expression of the target polynucleotide and compounds that specifically bind to or modulate the activity of the polypeptide  
XX  
SQ Sequence 1418 BP; 168 A; 542 C; 442 G; 266 T; 0 U; 0 Other;  
Query Match 0.9%; Score 20; DB 10; Length 1418;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Human secreted protein SECP-6 encoding cDNA SEQ ID NO:86.  
Human; secreted protein; SECP; cytostatic; antiarteriosclerotic; anticonvulsant; neurotropic; neuroprotective; cerebroprotective; anti-HIV; anti-allergic; anti-inflammatory; thymimetic; gene therapy; cancer; cell proliferative disorder; atherosclerosis; neurological disorder; epilepsy; Huntington's disease; stroke; immune disorder; allergy; AIDS; inflammatory disorder; developmental disorder; hypothyroidism; Cushing's syndrome; infection; gene; ss.  
Homo sapiens.  
WO2003016506-A2.  
27-FEB-2003.  
15-AUG-2002; 2002WO-US027143.  
17-AUG-2001; 2001US-0313249P.  
24-AUG-2001; 2001US-0314752P.  
07-SEP-2001; 2001US-0317818P.  
07-SEP-2001; 2001US-0317824P.  
21-SEP-2001; 2001US-0324040P.  
24-SEP-2001; 2001US-0324586P.  
02-NOV-2001; 2001US-0343980P.  
28-NOV-2001; 2001US-0334229P.  
13-FEB-2002; 2002US-0357002P.  
06-MAR-2002; 2002US-0362439P.  
19-MAR-2002; 2002US-0366041P.  
30-APR-2002; 2002US-0376988P.  
(INCY-) INCYTE GENOMICS INC.  
Tang TY, Warren BA, Gietzen KJ, Lal PG, Yue H, Honchell CD; Lehr-Mason PM, Burford N, Xu Y, Baughn MR, Duggan BM, Tran UK; Lee EA, Forsythe IJ, Richardson TW, Lee S, Thangavelu K, Yue H; Emerling BM, Walla NK, Azimzai Y, Sanjanwala B, Hafalia AJA; PI Borowsky ML, Nguyen DB, Ison CH, Astronoff A, Ding L, Lee SY; Becha SD, Ramkumar J, Gandhi AR, Jin P, Fu GK, Swarnakar A;  
WPI: 2003-278569/27.  
P-PSDB; ABR56731.  
New human secreted proteins (SECP), useful for diagnosing, treating and preventing diseases or conditions associated with the aberrant SECP expression e.g. cancer, AIDS, atherosclerosis, Huntington's disease, stroke, infections.  
Claim 5; Page 246; 286pp; English.  
ACC79026 to ACC79105 encode the human secreted proteins (I) given in ABR56726 to ABR56805, designated SECP-1 to SECP-80. SECP sequences can have cytostatic, antiarteriosclerotic, anticonvulsant, anti-inflammatory, neurotropic, neuroprotective, cerebroprotective, anti-HIV, anti-allergic and anticonvulsant activities, and can be used in gene therapy. The SECP proteins and polynucleotides can be used in diagnosing, treating and preventing diseases or conditions associated with the decreased expression or overexpression of SECP, such as cell proliferative (e.g. cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's disease, stroke), immune/inflammatory (e.g. AIDS, allergies) and developmental (e.g. hypothyroidism, Cushing's syndrome) disorders, or infections. They are also useful in assessing the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of SECP. The SECP or its fragments are useful in screening compounds for effectiveness as agonist or antagonist of the polypeptides, or in altering the expression of the target polynucleotide and compounds that specifically bind to or modulate the activity of the polypeptide  
Sequence 1418 BP; 168 A; 542 C; 442 G; 266 T; 0 U; 0 Other;  
Query Match 0.9%; Score 20; DB 10; Length 1418;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 GCAACAGCAGCAGCAGCGGC 62  
 |||||  
 Db 138 GCAACAGCAGCAGCAGCGGC 119

RESULT 69  
 ABL04193  
 ID ABL04193 standard; cDNA; 1565 BP.  
 XX  
 AC ABL04193;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 7061.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ss.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 XX 23-MAR-2001; 2001WO-US009231.  
 XX  
 XX 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 XX (PEKE ) PE CORP NY.  
 PA  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 DR  
 DR WPI; 2001-658860/75.  
 DR P-PSDB; ABB60090.  
 XX  
 XX New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.  
 XX  
 PS Claim 1; SEQ ID NO 7061; 21bp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 1565 BP; 397 A; 491 C; 416 G; 261 T; 0 U; 0 Other;

Query Match 0.9%; Score 20; DB 4; Length 1565;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 GCAACAGCAGCAGCAGCGGC 62  
 |||||  
 Db 234 GCAACAGCAGCAGCAGCGGC 253

RESULT 70  
 ADM47862  
 ID ADM47862 standard; DNA; 1891 BP.  
 XX  
 AC ADM47862;  
 XX  
 DT 03-JUN-2004 (first entry)

XX  
 DE Polynucleotide sequence #280 useful in producing transgenic plants.  
 XX  
 KW Plant; maize; rice; sorghum; thale cress; soybean; wheat; transgenic;  
 KW osmotic stress; sugar transport; cell cycle pathway; plant height;  
 KW carbohydrate transport; crop productivity; plant growth;  
 KW stress resistance; disease resistance; insect resistance; heat tolerance;  
 KW nitrogen assimilation; water stress tolerance;  
 KW photosynthetic carbon fixation; virus resistance; gene therapy; ds.  
 XX  
 OS Oryza sativa.  
 XX  
 AC US2003233670-A1.  
 PN  
 XX 18-DEC-2003.  
 PD  
 XX 04-DEC-2002; 2002US-00310154.  
 PF  
 XX 04-DEC-2001; 2001US-0337358P.  
 PR  
 XX (EDGE/) EDGERTON M D.  
 PA (CHOM/) CHOMET P S.  
 PA (LACC/) LACCETTI L B.  
 XX  
 XX Edgerton MD, Chomet PS, Laccetti LB;  
 PI  
 DR WPI; 2004-061374/06.  
 DR P-PSDB; ADM48230.  
 XX  
 XX New polynucleotide, useful for manipulating plant protein quality,  
 PT improving plant growth, yield and crop productivity or grain composition  
 PT or producing plants with improved properties.  
 PT  
 XX  
 PS Claim 1; SEQ ID NO 280; 144pp; English.  
 XX  
 CC The present invention relates to polynucleotide sequences, and the  
 CC proteins they encode. The sequences are isolated from a variety of  
 CC organisms such as plants (e.g. maize, rice, sorghum, thale cress,  
 CC soybean, and wheat), cyanobacteria, bacteria, yeast and other fungi. The  
 CC polynucleotide and polypeptide sequences of the invention are useful in  
 CC the production of transgenic plants that have improved properties. Also  
 CC disclosed are methods of producing fertile transgenic plants, preferably  
 CC maize, with desired phenotypes. The polynucleotide and polypeptide  
 CC sequences are useful for improving plants by providing protection against  
 CC osmotic stress, improving altering sugar transport and/or metabolism,  
 CC modifying the cell cycle pathway, reducing plant height, modifying  
 CC carbohydrate transport, improving crop productivity, improving plant  
 CC growth and stress resistance, improving disease resistance, improving  
 CC insect resistance, improving cold or heat tolerance, improving nitrogen  
 CC assimilation, improving stalk strength, improving water stress tolerance,  
 CC improving photosynthetic carbon fixation, improving biotic and abiotic  
 CC stress resistance, improving resistance to oxidative stress, providing  
 CC increased vigour, reducing senescence, and conferring virus resistance.  
 CC The present sequence represents a polynucleotide sequence of the  
 CC invention. Note: The sequence data for this patent is not provided in the  
 CC printed specification but is obtained in electronic format from the USPTO  
 CC website at seqdata.uspto.gov.  
 XX  
 SQ Sequence 1891 BP; 379 A; 669 C; 494 G; 349 T; 0 U; 0 Other;

Query Match 0.9%; Score 20; DB 12; Length 1891;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GCGGCAACAGCAGCAGCAGC 59  
 |||||  
 Db 1125 GCGGCAACAGCAGCAGCAGC 1144

RESULT 71  
 ADA71941  
 ID ADA71941 standard; DNA; 2000 BP.  
 XX



```
DE Human cDNA sequence SEQ ID NO:18501.
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
KW Homo sapiens.
OS
XX EP1074617-A2.
XX
XX PD 07-FEB-2001.
XX
XX PF 28-JUL-2000; 2000EP-00116126.
XX
XX PR 29-JUL-1999; 99JP-00248036.
XX
XX PR 27-AUG-1999; 99JP-00300253.
XX
XX PR 11-JAN-2000; 2000JP-00118776.
XX
XX PR 02-MAY-2000; 2000JP-00183767.
XX
XX PR 09-JUN-2000; 2000JP-00241899.
XX
XX PA (HELI-) HELIX RES INST.
XX
XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX PR Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX length cDNAs defined in the specification, and for the detection and/or
XX diagnosis of the abnormality of the proteins encoded by the full-length
XX cDNAs.
XX
XX PS Claim 8; SEQ ID NO 18501; 2537pp + Sequence Listing; English.
XX
XX CC The present invention describes primer sets for synthesising 5602 full-
XX length cDNAs defined in the specification. Where a primer set comprises:
XX (a) an oligo-dr primer and an oligonucleotide complementary to the
XX complementary strand of a polynucleotide which comprises one of the 5602
XX nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in the
XX specification. The primer sets can be used in antisense therapy and in
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
XX represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX oligonucleotides, all of which are used in the exemplification of the
XX present invention
XX
XX SQ Sequence 2370 BP; 639 A; 459 C; 440 G; 832 T; 0 U; 0 Other;
Query Match 0.9%; Score 20; DB 4; Length 2370;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 972 GAACAAAACCAAGCTTAA 991
DB 1151 GAACAAAACCAAGCTTAA 1132
RESULT 74
AAS72135
ID AAS72135 standard; cDNA; 2370 BP.
XX
XX AC AAS72135;
XX
XX DT 13-FEB-2002 (first entry)
```

```
XX DNA encoding novel human diagnostic protein #7939.
DE
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX OS
XX WO200175067-A2.
XX
XX PD 11-OCT-2001.
XX
XX PF 30-MAR-2001; 2001WO-US008631.
XX
XX PR 31-MAR-2000; 2000US-00540217.
XX
XX PR 23-AUG-2000; 2000US-00649167.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX PI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX
XX DR P-PSDB; ABG07948.
XX
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX PS Claim 1; SEQ ID NO 7939; 103pp; English.
XX
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
XX coding sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 2370 BP; 770 A; 396 C; 532 G; 672 T; 0 U; 0 Other;
Query Match 0.9%; Score 20; DB 5; Length 2370;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 790 ATGAGGAACCAAGCAGTATAT 809
DB 2040 ATGAGGAACCAAGCAGTATAT 2059
RESULT 75
ADG31190/c
ID ADG31190 standard; DNA; 2803 BP.
XX
XX AC ADG31190;
XX
XX DT 26-FEB-2004 (first entry)
XX
XX DT Novel mouse gene #15.
```

XX novel protein; enzyme; blood clotting disorder; circulatory disorder;  
KW mouse; murine; gene; ds.  
XX  
OS Mus musculus.  
XX WO2003089644-A1.  
XX  
XX 30-OCT-2003.  
XX  
XX 18-APR-2003; 2003WO-JP004983.  
XX  
XX 19-APR-2002; 2002JP-00117794.  
XX 26-APR-2002; 2002JP-00126250.  
XX 30-APR-2002; 2002JP-00128657.  
XX 02-MAY-2002; 2002JP-00130702.  
XX 02-MAY-2002; 2002JP-00130871.  
XX 02-MAY-2002; 2002JP-00130918.  
XX 02-MAY-2002; 2002JP-00130925.  
XX 04-DEC-2002; 2002JP-00352326.  
XX 04-DEC-2002; 2002JP-00352381.  
XX 04-DEC-2002; 2002JP-00352559.  
XX 04-DEC-2002; 2002JP-00352694.  
XX 04-DEC-2002; 2002JP-00352786.  
XX 04-DEC-2002; 2002JP-00352808.  
XX  
XX (RIKE ) RIKEN KK.  
XX (DNAP-) DNAPFORM KK.  
XX (MITU ) MITSUBISHI CHEM CORP.  
XX  
XX Hayashizaki Y, Kamiya M, Kubodera H;  
XX  
XX WPI; 2004-034530/03.  
XX P-PSDB; ADG31214.  
XX  
XX Protease inhibitors, proteases, phosphodiesterases, transferases,  
XX isomerases and substances that affect their activity or expression useful  
XX for treating associated disorders.  
XX  
XX Claim 13; SEQ ID NO 15; 524pp; Japanese.  
XX  
XX The invention comprises the amino acid and coding sequences of proteins  
XX which have enzyme activity or protease, phosphodiesterase, transferase,  
XX or isomerase inhibitor activity. The DNA and protein sequences of the  
XX invention are useful for treating disorders associated with the proteins,  
XX such as blood clotting or circulatory disorders. The present DNA sequence  
XX represents a novel mouse gene of the invention.  
XX  
XX Sequence 2803 BP; 546 A; 848 C; 854 G; 555 T; 0 U; 0 Other;  
SQ

Query Match 0.9%; Score 20; DB 12; Length 2803;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 GCAACAGCAGCAGCAGCGGC 62  
Db 338 GCAACAGCAGCAGCAGCGGC 319

Search completed: March 25, 2005, 17:53:06  
Job time : 1171 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 20:35:20 ; Search time 1231 Seconds  
(without alignments)  
11250.599 Million cell updates/sec

Title: US-09-436-184-3  
Perfect score: 2324  
Sequence: 1 cggacgtgcaatggccag.....cttgggaactctggagaga 2324

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 5552208 seqs, 2979665951 residues

Word size : 0

Total number of hits satisfying chosen parameters: 11104416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Published Applications NA.\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/ECT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/ECTUS\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2324	100.0	2324	9 US-09-903-248-3	Sequence 3, Appli
2	2324	100.0	2324	9 US-09-859-604-3	Sequence 3, Appli
3	2324	100.0	2324	9 US-09-903-063-3	Sequence 3, Appli
4	2324	100.0	2324	9 US-09-903-216-3	Sequence 3, Appli
5	2324	100.0	2324	9 US-09-903-199-3	Sequence 3, Appli
6	2324	100.0	2324	9 US-09-903-023-3	Sequence 3, Appli
7	2324	100.0	2324	10 US-09-436-184-3	Sequence 3, Appli
8	2171	93.4	5358	15 US-10-084-817-63	Sequence 63, Appli
9	895	38.5	2680	16 US-10-096-534-37	Sequence 37, Appli
10	895	38.5	2680	16 US-10-263-909-3	Sequence 3, Appli
11	640	27.5	2442	9 US-09-736-457-1791	Sequence 1791, Ap

12	640	27.5	2442	9 US-09-902-941-1791	Sequence 1791, Ap
13	640	27.5	2442	9 US-09-849-628-1791	Sequence 1791, Ap
14	640	27.5	2442	14 US-10-017-754-1791	Sequence 1791, Ap
15	640	27.5	2442	16 US-10-113-872-1791	Sequence 1791, Ap
16	640	27.5	2442	17 US-10-283-017-1791	Sequence 1791, Ap
17	615	26.5	3110	14 US-10-198-846-9938	Sequence 9938, Ap
18	459	19.8	502	10 US-09-918-995-34279	Sequence 34279, A
19	359	15.4	660	11 US-09-969-034-3202	Sequence 3202, Ap
20	315	13.6	366	17 US-10-242-535A-25661	Sequence 25661, A
21	315	13.6	366	17 US-10-085-783A-25661	Sequence 25661, A
22	249	10.7	495	14 US-10-066-543-277	Sequence 277, App
23	135	5.8	191	17 US-10-243-535A-2707	Sequence 2707, Ap
24	135	5.8	191	17 US-10-085-783A-2707	Sequence 2707, Ap
25	104	4.5	161	16 US-10-029-386-26964	Sequence 26964, A
26	104	4.5	592	16 US-10-029-386-13264	Sequence 13264, A
27	78	3.4	450	13 US-10-027-632-2825	Sequence 2825, Ap
28	78	3.4	450	17 US-10-027-632-2825	Sequence 2825, Ap
29	72	3.1	1382	17 US-10-443-622-24	Sequence 24, Appli
30	60	2.6	60	10 US-09-908-975-11070	Sequence 11070, A
31	59	2.5	2208	17 US-10-116-275-322	Sequence 322, App
32	59	2.5	2208	18 US-10-755-889-177	Sequence 177, App
33	51	2.2	583	11 US-09-969-034-3266	Sequence 3266, Ap
34	44	1.9	591	13 US-10-027-632-7020	Sequence 7020, Ap
35	44	1.9	591	17 US-10-027-632-7020	Sequence 7020, Ap
36	39	1.7	453	13 US-10-027-632-56205	Sequence 56205, A
37	39	1.7	453	13 US-10-027-632-310460	Sequence 310460, A
38	39	1.7	453	17 US-10-027-632-56205	Sequence 56205, A
39	39	1.7	453	17 US-10-027-632-310460	Sequence 310460, A
40	24	1.0	859	17 US-10-424-599-97119	Sequence 97119, A
41	23	1.0	2135	18 US-10-425-115-124491	Sequence 124491, A
42	21	0.9	681	18 US-10-437-963-73067	Sequence 73067, A
43	21	0.9	681	18 US-10-425-115-42949	Sequence 42949, A
44	21	0.9	1845	16 US-10-308-503-188	Sequence 188, App
45	20	0.9	20	9 US-09-859-604-10	Sequence 10, Appli
46	20	0.9	20	9 US-09-859-604-11	Sequence 11, Appli
47	20	0.9	20	9 US-09-859-604-12	Sequence 12, Appli
48	20	0.9	183	18 US-10-425-115-21348	Sequence 21348, A
49	20	0.9	432	18 US-10-425-115-22449	Sequence 22449, A
50	20	0.9	489	18 US-10-767-701-26002	Sequence 26002, A
51	20	0.9	495	17 US-10-424-599-89196	Sequence 89196, A
52	20	0.9	550	16 US-10-029-386-9144	Sequence 9144, Ap
53	20	0.9	567	13 US-10-027-632-82527	Sequence 82527, A
54	20	0.9	567	13 US-10-027-632-322022	Sequence 322022, A
55	20	0.9	567	17 US-10-027-632-82527	Sequence 82527, A
56	20	0.9	567	17 US-10-027-632-322022	Sequence 322022, A
57	20	0.9	567	17 US-10-027-632-322023	Sequence 322023, A
58	20	0.9	750	18 US-10-425-115-112661	Sequence 112661, A
59	20	0.9	771	18 US-10-437-963-71261	Sequence 71261, A
60	20	0.9	804	18 US-10-822-613-1	Sequence 1, Appli
61	20	0.9	843	18 US-10-437-963-24406	Sequence 24406, A
62	20	0.9	924	18 US-10-437-963-51503	Sequence 51503, A
63	20	0.9	936	18 US-10-646-620A-8	Sequence 8, Appli
64	20	0.9	1011	18 US-10-425-115-180542	Sequence 180542, A
65	20	0.9	1072	9 US-09-853-386-105	Sequence 105, App
66	20	0.9	1109	10 US-09-397-945-80	Sequence 80, Appli
67	20	0.9	1109	17 US-10-653-595-80	Sequence 80, Appli
68	20	0.9	1235	9 US-09-960-288-1	Sequence 1, Appli
69	20	0.9	1245	17 US-10-287-971-37	Sequence 37, Appli
70	20	0.9	1282	17 US-10-425-114-23644	Sequence 23644, A
71	20	0.9	1435	18 US-10-425-115-19062	Sequence 19062, A
72	20	0.9	1439	19 US-10-913-196-2	Sequence 2, Appli
73	20	0.9	1850	17 US-10-425-114-34740	Sequence 34740, A
74	20	0.9	1891	17 US-10-310-154-280	Sequence 280, App
75	20	0.9	1895	18 US-10-425-115-16475	Sequence 16475, A
76	20	0.9	2000	17 US-10-260-238-2021	Sequence 2021, Ap
77	20	0.9	8772	9 US-09-788-711A-3	Sequence 3, Appli
78	20	0.9	8871	9 US-09-788-711A-1	Sequence 1, Appli
79	20	0.9	9321	17 US-10-311-623-21	Sequence 21, Appli
80	20	0.9	10531	15 US-10-225-567A-523	Sequence 523, App
81	20	0.9	24370	17 US-10-292-798-931	Sequence 931, App
82	20	0.9	47777	19 US-10-741-600-17649	Sequence 17649, A
83	20	0.9	52677	13 US-10-087-192-130	Sequence 130, App



c 85 20 0.9 159400 18 US-10-450-826-33 Sequence 33, Appl  
86 20 0.9 822900 17 US-10-292-798-1393 Sequence 1393, Ap  
c 87 20 0.9 2731748 18 US-10-297-465A-1 Sequence 1, Appli  
c 88 19 0.8 193 18 US-10-437-963-88593 Sequence 88593, A  
c 89 19 0.8 233 17 US-10-424-599-1798 Sequence 1798, Ap  
c 90 19 0.8 347 18 US-10-425-115-14296 Sequence 14296, A  
c 91 19 0.8 353 9 US-09-783-590-4838 Sequence 4838, Ap  
c 92 19 0.8 389 9 US-09-960-352-868 Sequence 868, App  
c 93 19 0.8 394 9 US-09-960-352-3447 Sequence 3447, Ap  
c 94 19 0.8 400 9 US-09-960-352-7474 Sequence 7474, Ap  
c 95 19 0.8 404 9 US-09-960-352-8848 Sequence 8848, Ap  
c 96 19 0.8 483 13 US-10-062-254-197 Sequence 197, App  
c 97 19 0.8 497 17 US-10-425-114-29177 Sequence 29177, A  
c 98 19 0.8 500 17 US-10-621-901-185 Sequence 185, App  
c 99 19 0.8 505 18 US-10-437-963-85957 Sequence 85957, A  
c 100 19 0.8 507 10 US-09-918-995-15209 Sequence 15209, A

## ALIGNMENTS

## RESULT 1

US-09-903-248-3

; Sequence 3, Application US/09903248

; Patent No. US20020102263A1

; GENERAL INFORMATION:

; APPLICANT: Wands, Jack R.

; APPLICANT: de la Monte, Suzanne M.

; APPLICANT: Ince, Nedim

; APPLICANT: Carlson, Rolf I.

; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS

; FILE REFERENCE: 21486-032 DIV5

; CURRENT APPLICATION NUMBER: US/09/903,248

; CURRENT FILING DATE: 2001-07-11

; PRIOR APPLICATION NUMBER: 09/436,184

; PRIOR FILING DATE: 1999-11-08

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 2324

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-09-903-248-3

Query Match 100.0%; Score 2324; DB 9; Length 2324;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CGGACCGTGC	AAATGGCC	CAGCGT	GAAGAT	GC	CAAGAG	CAGCGG	CAACAG	CAGCAG	CAGCAG	CAGCAG	60
Db	1	CGGACCGTGC	AAATGGCC	CAGCGT	GAAGAT	GC	CAAGAG	CAGCGG	CAACAG	CAGCAG	CAGCAG	CAGCAG	60
Qy	61	GCTCGG	CAGCGT	AGCAGT	CGGG	CAGCAG	CAGCGG	CGGG	CGGG	CGGG	CGGG	CGGG	120
Db	61	GCTCGG	CAGCGT	AGCAGT	CGGG	CAGCAG	CAGCGG	CGGG	CGGG	CGGG	CGGG	CGGG	120
Qy	121	AGCATG	AGGACACA	GAAGAT	GGGAGG	AAAGCG	GACTCT	CGGGA	AACTTCA	TTCTT	CACTT	CACTT	180
Db	121	AGCATG	AGGAGACA	CAAGAT	GGGAGG	AAAGCG	GACTCT	CGGGA	AACTTCA	TTCTT	CACTT	CACTT	180
Qy	181	GGTTAT	TGTTG	CAATG	CTG	TGCG	CGTCT	CGA	CACTG	TAGT	GTG	TTGGT	240
Db	181	GGTTAT	TGTTG	CAATG	CTG	TGCG	CGTCT	CGA	CACTG	TAGT	GTG	TTGGT	240
Qy	241	TTGTTG	ACTAT	GAGG	AGTTCT	TAGG	AACTAG	GAATCT	TATG	ATG	CTG	ATG	300
Db	241	TTGTTG	ACTAT	GAGG	AGTTCT	TAGG	AACTAG	GAATCT	TATG	ATG	CTG	ATG	300
Qy	301	ATTTT	TGATG	TGATG	ATG	CCCAAG	TTTTT	TAT	TAGG	ACTT	TAA	GAGAG	360
Db	301	ATTTT	TGATG	TGATG	ATG	CCCAAG	TTTTT	TAT	TAGG	ACTT	TAA	GAGAG	360
Qy	361	CAGCAG	TCCC	CCCA	AGAG	CGT	GAG	CCCA	CAC	ACT	GAG	CCCG	420
Db	361	CAGCAG	TCCC	CCCA	AGAG	CGT	GAG	CCCA	CAC	ACT	GAG	CCCG	420

Db 1441 ATGAGAGGTCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTATGGCTTCA 1500  
Qy 1501 TCTGAGGCGCAGAACAAATTCCTGAGAGCATCCCATATTTAAAGGAAGGATGAAT 1560  
Db 1501 TCTGAGGCGCAGAACAAATTCCTGAGAGCATCCCATATTTAAAGGAAGGATGAAT 1560  
Qy 1561 CCGGAGATCTGGCACTGATGATGGGAGATTTTATTTCCACCTGGGGGATGCCATGCGA 1620  
Db 1561 CCGGAGATCTGGCACTGATGATGGGAGATTTTATTTCCACCTGGGGGATGCCATGCGA 1620  
Qy 1621 GGGTTGGGAAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAAGAGAGGACATTTG 1680  
Db 1621 GGGTTGGGAAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAAGAGAGGACATTTG 1680  
Qy 1681 CATCTGCTGGCAACGGCTCACTCTACATGTGAATGGACTGAAGCAGCAGCTTGTGGA 1740  
Db 1681 CATCTGCTGGCAACGGCTCACTCTACATGTGAATGGACTGAAGCAGCAGCTTGTGGA 1740  
Qy 1741 CCCGAAAGAAACCGGCTACACAGATTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
Db 1741 CCCGAAAGAAACCGGCTACACAGATTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
Qy 1801 TCCGAGATGAAGGCTTGCAGTGTATGATGATAAAGCCAAAGGCTCTTCTGCTGAGGATG 1860  
Db 1801 TCCGAGATGAAGGCTTGCAGTGTATGATGATAAAGCCAAAGGCTCTTCTGCTGAGGATG 1860  
Qy 1861 AAAACCTGAGGAAAGAGGAGCTGGAGCCAGTTCACGCTGTGGCAGCAGGAAGAA 1920  
Db 1861 AAAACCTGAGGAAAGAGGAGCTGGAGCCAGTTCACGCTGTGGCAGCAGGAAGAA 1920  
Qy 1921 ATGAAATGCTGCAAGAGGAGCTCTTAAACCTGTACCTTACTAGAAAGTCTCCCGAGA 1980  
Db 1921 ATGAAATGCTGCAAGAGGAGCTCTTAAACCTGTACCTTACTAGAAAGTCTCCCGAGA 1980  
Qy 1981 CAACAGGATGAGAGGAGGACAGATCAAAATATTCATCATGACCCCGGAGCTCACGTGT 2040  
Db 1981 CAACAGGATGAGAGGAGGACAGATCAAAATATTCATCATGACCCCGGAGCTCACGTGT 2040  
Qy 2041 GGGCGCACACAGGGCCACACAACTGAGGCTCCGAATGACCTGGGCTTGGTATTCCTCA 2100  
Db 2041 GGGCGCACACAGGGCCACACAACTGAGGCTCCGAATGACCTGGGCTTGGTATTCCTCA 2100  
Qy 2101 AGGAAGGCTGACAGATTCGATGTCGCAAGCAGCAGCAGCTGGGAGGAGGAGGAGTGC 2160  
Db 2101 AGGAAGGCTGACAGATTCGATGTCGCAAGCAGCAGCAGCTGGGAGGAGGAGGAGTGC 2160  
Qy 2161 TCATCTTTGATGACTCTTTGAGCAGAGGTATGGCAGGATGCTCATCTTCCGCTGA 2220  
Db 2161 TCATCTTTGATGACTCTTTGAGCAGAGGTATGGCAGGATGCTCATCTTCCGCTGA 2220  
Qy 2221 TATTATCTGATGATGTTGGCATCCGAACTGACACACAGCAGAGCAGCAGCTTCCAG 2280  
Db 2221 TATTATCTGATGATGTTGGCATCCGAACTGACACACAGCAGAGCAGCAGCTTCCAG 2280  
Qy 2281 CAATTTAGCATGAATTCATGCAAGCTTTGGAAACTCTGGAGAGA 2324  
Db 2281 CAATTTAGCATGAATTCATGCAAGCTTTGGAAACTCTGGAGAGA 2324

## RESULT 2

US-09-859-604-3  
; Sequence 3, Application US/09859604  
; Patent No. US20020110559A1  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M  
; APPLICANT: Deutch, Alan H  
; APPLICANT: Ghanbari, Hossein A  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 CIP  
; CURRENT APPLICATION NUMBER: US/09/859,604  
; CURRENT FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 09/436,184

; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-859-604-3

Query Match 100.0%; Score 2324; DB 9; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CGGACCGTGCATATGGCCAGCGTAAAGATGCCAAGCAGCGGCAACAGCAGCAGCAGCG 60  
Db 1 CGGACCGTGCATATGGCCAGCGTAAAGATGCCAAGCAGCGGCAACAGCAGCAGCAGCG 60  
Qy 61 GCTCCGGCAGCGGTAGCAGAGTGGCGGCGCAGCAGCGCCCGGGCCCGGAGAGACAA 120  
Db 61 GCTCCGGCAGCGGTAGCAGAGTGGCGGCGCAGCAGCGCCCGGGCCCGGAGAGACAA 120  
Qy 121 AGCATGGAGGACACAAGAAATGGGAGGAAAGGCGGACTCTCGGAACTTCTATCTTCCAGT 180  
Db 121 AGCATGGAGGACACAAGAAATGGGAGGAAAGGCGGACTCTCGGAACTTCTATCTTCCAGT 180  
Qy 181 GGTATTATGGTGAATGCAATTCGTCGGCGCTCTGGAACATCTGTAGCTGTCTGTTTGTGATC 240  
Db 181 GGTATTATGGTGAATGCAATTCGTCGGCGCTCTGGAACATCTGTAGCTGTCTGTTTGTGATC 240  
Qy 241 TTGTTGACTATGAGGAAGTCTTAGGAAACTAGGAATCTATGATGCTGTGATGTTGATGAG 300  
Db 241 TTGTTGACTATGAGGAAGTCTTAGGAAACTAGGAATCTATGATGCTGTGATGTTGATGAG 300  
Qy 301 ATTTTGTATGTTGATGATGCAAAAGTTTATTAGGACTTAAAGAGAGATCTACTTTCAGAGC 360  
Db 301 ATTTTGTATGTTGATGATGCAAAAGTTTATTAGGACTTAAAGAGAGATCTACTTTCAGAGC 360  
Qy 361 CAGCAGTCCCGCAGAGAGGCTGACCACTGTAGCCCGGAGGAGGAGTCTCTGTGG 420  
Db 361 CAGCAGTCCCGCAGAGAGGCTGACCACTGTAGCCCGGAGGAGGAGTCTCTGTGG 420  
Qy 421 AGGCAAGAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCCATG 480  
Db 421 AGGCAAGAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCCATG 480  
Qy 481 AAATGGTACACGCAAGAACATGTTGAGGAGGAGAACTTTCGCAACAGAGAGATGGACCCACAG 540  
Db 481 AAATGGTACACGCAAGAACATGTTGAGGAGGAGAACTTTCGCAACAGAGAGATGGACCCACAG 540  
Qy 541 GAGAACCAACAAGAGGATGATGATTTCTTATGGCAGCTGATGTAGATGATGATGATG 600  
Db 541 GAGAACCAACAAGAGGATGATGATTTCTTATGGCAGCTGATGTAGATGATGATGATG 600  
Qy 601 AGACCTTGGAACTCTGAACTATCTCATGAAGAAACCGAGCATAGTTACCACTGGGAGAGA 660  
Db 601 AGACCTTGGAACTCTGAACTATCTCATGAAGAAACCGAGCATAGTTACCACTGGGAGAGA 660  
Qy 661 CAGTTTCAAGACTGTATATCAGGATATGGAAGATGATGTCTGAGCAGGAGAAATCCAG 720  
Db 661 CAGTTTCAAGACTGTATATCAGGATATGGAAGATGATGTCTGAGCAGGAGAAATCCAG 720  
Qy 721 ATTCCAGTGAACCCAGTAGTAGAAGATGAAGATTCACCATGATACAGATGATGTAAAT 780  
Db 721 ATTCCAGTGAACCCAGTAGTAGAAGATGAAGATTCACCATGATACAGATGATGTAAAT 780  
Qy 781 ACCAAGTCTATGAGGAAACACGACTATATGAACTCTAGAAATGAGGAGATGAAATCA 840  
Db 781 ACCAAGTCTATGAGGAAACACGACTATATGAACTCTAGAAATGAGGAGATGAAATCA 840  
Qy 841 CAGAGTAACTGCTCCCTCCCTGAGGATATCTCTGTAGAGATTCACAGGTAAATTTAGAG 900  
Db 841 CAGAGTAACTGCTCCCTCCCTGAGGATATCTCTGTAGAGATTCACAGGTAAATTTAGAG 900

QY 901 AAGTAAGCATTTTCTCTGGAAGAACAGCAGAGGAGTACACAGAGAAACAAATAGAAAA 960  
DB 901 AAGTAAGCATTTTCTCTGGAAGAACAGCAGAGGAGTACACAGAGAAACAAATAGAAAA 960  
QY 961 CAGATGATCCAGAACAAAAAGCTTAAAGAAAAAGAGCCCTTAAATTTTAAATTAAT 1020  
DB 961 CAGATGATCCAGAACAAAAAGCTTAAAGAAAAAGAGCCCTTAAATTTTAAATTAAT 1020  
QY 1021 TTGTAAGACTATTAAGAGCTGAACTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1080  
DB 1021 TTGTAAGACTATTAAGAGCTGAACTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1080  
QY 1081 TTGAGGAGCAGTGAATGCAATTAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140  
DB 1081 TTGAGGAGCAGTGAATGCAATTAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140  
QY 1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGGAGGAGTAATGAGG 1200  
DB 1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGGAGGAGTAATGAGG 1200  
QY 1201 TGCTACGTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCTACCTGATGTCCTCGAG 1260  
DB 1201 TGCTACGTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCTACCTGATGTCCTCGAG 1260  
QY 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTCTGCTCAGACAGGCAACAAATTTCTAGGTCAATGA 1320  
DB 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTCTGCTCAGACAGGCAACAAATTTCTAGGTCAATGA 1320  
QY 1321 GAGGTTCCCTGCTTACCTCGAGAGATTAATTAATTTTCCCAATGATACCTTCTTAA 1380  
DB 1321 GAGGTTCCCTGCTTACCTCGAGAGATTAATTAATTTTCCCAATGATACCTTCTTAA 1380  
QY 1381 AAAATGACCTTGGCTGGGATACCTCTGATAGGAGATAATGACAAATGCAAGAAAGTTT 1440  
DB 1381 AAAATGACCTTGGCTGGGATACCTCTGATAGGAGATAATGACAAATGCAAGAAAGTTT 1440  
QY 1441 ATGAAGAGGTGCTGAGTGTGACACCTAATGATGCTTTGCTAAAGTCCATTTATGGCTTCA 1500  
DB 1441 ATGAAGAGGTGCTGAGTGTGACACCTAATGATGCTTTGCTAAAGTCCATTTATGGCTTCA 1500  
QY 1501 TCCTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCATTTTAAAGGAAGGAATAGAT 1560  
DB 1501 TCCTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCATTTTAAAGGAAGGAATAGAT 1560  
QY 1561 CCGGAGATCTCGCACTGATGATGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGA 1620  
DB 1561 CCGGAGATCTCGCACTGATGATGGAGATTTTATTTCCACCTGGGGGATGCCATGCAGA 1620  
QY 1621 GGGTTGGGAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAAGAGAGGACACTTTG 1680  
DB 1621 GGGTTGGGAACAAAGAGGCATATAAGTGGTATGAGCTTGGGCAAGAGAGGACACTTTG 1680  
QY 1681 CATCTGTCTGGCAACGCTTCACTCAATGTAAGTGAAGTGAAGCAAGAGGAGGAGTGA 1740  
DB 1681 CATCTGTCTGGCAACGCTTCACTCAATGTAAGTGAAGTGAAGCAAGAGGAGGAGTGA 1740  
QY 1741 CCCCCAAAGAAACGGGCTTACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
DB 1741 CCCCCAAAGAAACGGGCTTACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
QY 1801 TCCGAGATGAAGCCCTTGCAGTGTGATGAATAAGCCAAAGGCTCTCTCTGCTGAGGATG 1860  
DB 1801 TCCGAGATGAAGCCCTTGCAGTGTGATGAATAAGCCAAAGGCTCTCTCTGCTGAGGATG 1860  
QY 1861 AAAACCTGAGGGAAGAAAGGGGACTGAGGCGAGTTACGCTGTGGCAGCAAGGAAAGAA 1920  
DB 1861 AAAACCTGAGGGAAGAAAGGGGACTGAGGCGAGTTACGCTGTGGCAGCAAGGAAAGAA 1920  
QY 1921 ATGAAATGCTGCAAGAGGCTCTTAAACCTGTATACCTTCTAGAAAGTTTCCCGAGA 1980  
DB 1921 ATGAAATGCTGCAAGAGGCTCTTAAACCTGTATACCTTCTAGAAAGTTTCCCGAGA 1980  
QY 1981 CAACAGGATGCAGAGAGGACAGATCAAAATATTCATCATGACCCCGGACTCAGGTGT 2040

DB 1981 CAACAGGATGCAGAGAGGACAGATCAAAATATTCATCATGCAACCCCGGACTCAGGTGT 2040  
QY 2041 GCGCGCACACAGGCGCCCAAAATCTGAGGCTCCGAATGCACTGGGCTTGGTATCCCA 2100  
DB 2041 GCGCGCACACAGGCGCCCAAAATCTGAGGCTCCGAATGCACTGGGCTTGGTATCCCA 2100  
QY 2101 AGGAAGCTTCAAGATTTCCGATGTGCAACAGAGACCAAGGACCTGGGAGGAGCAAGGTGC 2160  
DB 2101 AGGAAGCTTCAAGATTTCCGATGTGCAACAGAGACCAAGGACCTGGGAGGAGCAAGGTGC 2160  
QY 2161 TCATCTTTGATGATCTCTTTGAGCAGAGGTATGCGAGATGCTTATCTTTCCGCTGA 2220  
DB 2161 TCATCTTTGATGATCTCTTTGAGCAGAGGTATGCGAGATGCTTATCTTTCCGCTGA 2220  
QY 2221 TATTCATCTGATGATGTGTCATCTCCGAATCTGACACCAAGAGACGACGACCTTTCCAG 2280  
DB 2221 TATTCATCTGATGATGTGTCATCTCCGAATCTGACACCAAGAGACGACGACGACCTTTCCAG 2280  
QY 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2324  
DB 2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2324

RESULT 3  
US-09-903-063-3  
; Sequence 3, Application US/09903063  
; Patent No. US20020114810A1  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R., Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV3  
; CURRENT APPLICATION NUMBER: US/09/903,063  
; CURRENT FILING DATE: 2001-10-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-903-063-3

Query Match 100.0%; Score 2324; DB 9; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGACCCGTGCAATGGCCAGCGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
DB 1 CCGACCCGTGCAATGGCCAGCGTAAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
QY 61 GCTCCGCGAGCGGTAGCAGCGAGTGCAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 120  
DB 61 GCTCCGCGAGCGGTAGCAGCGAGTGCAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 120  
QY 121 AGCATGGAGGACACAAGAAATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
DB 121 AGCATGGAGGACACAAGAAATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180  
QY 181 GGTTCATGCTGATTTGCAATGCTGGCGCTGCGAATCTGAGCTGCTGTTGGTTTGCATC 240  
DB 181 GGTTCATGCTGATTTGCAATGCTGGCGCTGCGAATCTGAGCTGCTGTTGGTTTGCATC 240  
QY 241 TTGTTGATCTATGAGAGAGTTCTAGGAAATCTAGGAAATCTATGATGCTGATGATGAG 300  
DB 241 TTGTTGATCTATGAGAGAGTTCTAGGAAATCTAGGAAATCTATGATGCTGATGATGAG 300  
QY 301 ATTTTGTGCTGATGATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360

Db 301 ATTTTGAATGTGATGATCCCAAGTTTATTTATGAGCTTAAAGAGAGATCTACTTCAGAGC 360  
Qy 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGAGAGAGGTTCTCTGTGG 420  
Db 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGAGAGGTTCTCTGTGG 420  
Qy 421 AGCAGAAACCCAGAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCCTCATG 480  
Db 421 AGCAGAAACCCAGAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCCTCATG 480  
Qy 481 AAATGGTACAGCAGAAACATGTTGAGGGGAGAACTTGCAACAAAGAGATGGACCCACAG 540  
Db 481 AAATGGTACAGCAGAAACATGTTGAGGGGAGAACTTGCAACAAAGAGATGGACCCACAG 540  
Qy 541 GAGAACCAACAAAGAGGATGATGAGTTTCTTATGGGACTGATGTAGATGATGATTTG 600  
Db 541 GAGAACCAACAAAGAGGATGATGAGTTTCTTATGGGACTGATGTAGATGATGATTTG 600  
Qy 601 AGACCTCGAACCTGAGATATCTCATGAAGAACCGAGCATAGTTTACCAAGTGAAGAGA 660  
Db 601 AGACCTCGAACCTGAGATATCTCATGAAGAACCGAGCATAGTTTACCAAGTGAAGAGA 660  
Qy 661 CAGTTTCAACAGATGTAATCAGGATATGGAAGAGATGATGTTCTGAGCAGCAAAATCCAG 720  
Db 661 CAGTTTCAACAGATGTAATCAGGATATGGAAGAGATGATGTTCTGAGCAGCAAAATCCAG 720  
Qy 721 ATTCCAGTGAACAGTATGAGATGAAGATTTGACCATGATACAGATGATGATTAACAT 780  
Db 721 ATTCCAGTGAACAGTATGAGATGAAGATTTGACCATGATACAGATGATGATTAACAT 780  
Qy 781 ACCAAGTCTATGAGAACACAGCAGTATGTAACCTCTGAAATGAGAGGATGAATCA 840  
Db 781 ACCAAGTCTATGAGAACACAGCAGTATGTAACCTCTGAAATGAGAGGATGAATCA 840  
Qy 841 CAGAGTAACTGCTCCCTCCTGAGGATTAATCCTGTAGAGATTTCAAGGATTTGTAGAAG 900  
Db 841 CAGAGTAACTGCTCCCTCCTGAGGATTAATCCTGTAGAGATTTCAAGGATTTGTAGAAG 900  
Qy 901 AAGTAAAGATTTTCTGTGGAGAAACAGCAGGAGATGACCCAGCAAAACAAATAGAAAAA 960  
Db 901 AAGTAAAGATTTTCTGTGGAGAAACAGCAGGAGATGACCCAGCAAAACAAATAGAAAAA 960  
Qy 961 CAGATGATCAGAAACAAAGAGTAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
Db 961 CAGATGATCAGAAACAAAGAGTAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020  
Qy 1021 TTGATAAGACTATTAAAGCTGAATGATGCTGAGAAACCTCGTAAAGGGGAAAAA 1080  
Db 1021 TTGATAAGACTATTAAAGCTGAATGATGCTGAGAAACCTCGTAAAGGGGAAAAA 1080  
Qy 1081 TTGAGGAAAGAGTGAATGCAATTTAAAGAACTAGTACGCAAAATCCCTCAGAGTCCACGAG 1140  
Db 1081 TTGAGGAAAGAGTGAATGCAATTTAAAGAACTAGTACGCAAAATCCCTCAGAGTCCACGAG 1140  
Qy 1141 CAAAGATATGGGAAAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGAGAGAGAG 1200  
Db 1141 CAAAGATATGGGAAAGGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAGAGAGAGAG 1200  
Qy 1201 TGCTACGTGGAGCCATCAGACCTACCAAGAGGTGGCCAGCTACCTGATGCTCCGTGAG 1260  
Db 1201 TGCTACGTGGAGCCATCAGACCTACCAAGAGGTGGCCAGCTACCTGATGCTCCGTGAG 1260  
Qy 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320  
Db 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320  
Qy 1321 GAGGTTCCCTGCTTACCTCCTCAGAGATTAAGTTTCAACTATTTCCCAATGATCTTCCCTAA 1380  
Db 1321 GAGGTTCCCTGCTTACCTCCTCAGAGATTAAGTTTCAACTATTTCCCAATGATCTTCCCTAA 1380  
Qy 1381 AAAATGACCTTGGGCTGGGATACCTCTGATAGGAGATTAATGCAATGCAAGAGAGAGAG 1440  
Db 1381 AAAATGACCTTGGGCTGGGATACCTCTGATAGGAGATTAATGCAATGCAAGAGAGAGAG 1440

Qy 1441 ATGAAGAGTCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTAATGGCTTCA 1500  
Db 1441 ATGAAGAGTCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTAATGGCTTCA 1500  
Qy 1501 TCTTGAAGGCAAGAAACAAATTTGCTGAGAGATCCCATATTTTAAAGAGAGAGATAGAAT 1560  
Db 1501 TCTTGAAGGCAAGAAACAAATTTGCTGAGAGATCCCATATTTTAAAGAGAGAGATAGAAT 1560  
Qy 1561 CCGAGATCTCGGCACTGATGATGGGAGATTTTATTTCCACTGGGGATGCCATGCGA 1620  
Db 1561 CCGAGATCTCGGCACTGATGATGGGAGATTTTATTTCCACTGGGGATGCCATGCGA 1620  
Qy 1621 GGGTTGGGAAACAAAGAGGCAATTAAGTGTATGAGCTTTGGGCAACAGAGAGGACACTTTG 1680  
Db 1621 GGGTTGGGAAACAAAGAGGCAATTAAGTGTATGAGCTTTGGGCAACAGAGAGGACACTTTG 1680  
Qy 1681 CATCTGTCTGCAACGCTCTCTACAATGTGAATGGAATGAAAGCAACGCTTTGGTGA 1740  
Db 1681 CATCTGTCTGCAACGCTCTCTACAATGTGAATGGAATGAAAGCAACGCTTTGGTGA 1740  
Qy 1741 CCCCAGAGAAACCGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAGAACTGGAAGTTAA 1800  
Db 1741 CCCCAGAGAAACCGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAGAACTGGAAGTTAA 1800  
Qy 1801 TCCGAGATGAAGGCTTTGCAAGTATGATGATAAGCCAAAGGCTCTCTCTGCTGAGGATG 1860  
Db 1801 TCCGAGATGAAGGCTTTGCAAGTATGATGATAAGCCAAAGGCTCTCTCTGCTGAGGATG 1860  
Qy 1861 AAAACCTGAGGAAAGAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAG 1920  
Db 1861 AAAACCTGAGGAAAGAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAGGGAAG 1920  
Qy 1921 ATGAAATGCTGCAAGAGGAGCTCTAAACCTGTACCTTACTAGAAAGAGTTCCCGAGA 1980  
Db 1921 ATGAAATGCTGCAAGAGGAGCTCTAAACCTGTACCTTACTAGAAAGAGTTCCCGAGA 1980  
Qy 1981 CAAAGAGATGAGAGAGAGAGATCAATATTTCCATCATGACCCCGGAGCTCACGTGT 2040  
Db 1981 CAAAGAGATGAGAGAGAGAGATCAATATTTCCATCATGACCCCGGAGCTCACGTGT 2040  
Qy 2041 GCGCGCACACAGGCGCCCAAACTGCAAGTCCGAGTCCGAGTCCGAGTCCGAGTCCCA 2100  
Db 2041 GCGCGCACACAGGCGCCCAAACTGCAAGTCCGAGTCCGAGTCCGAGTCCGAGTCCCA 2100  
Qy 2101 AGGAGGCTGCAAGATTTGATGTGCAACAGAGACCCAGAGAGAGAGAGAGAGAGAGAG 2160  
Db 2101 AGGAGGCTGCAAGATTTGATGTGCAACAGAGACCCAGAGAGAGAGAGAGAGAGAGAG 2160  
Qy 2161 TCATCTTTGATGACTCTTTTGAGCAGAGGATGATGAGGATGCTCATCTTTCCGCTGA 2220  
Db 2161 TCATCTTTGATGACTCTTTTGAGCAGAGGATGATGAGGATGCTCATCTTTCCGCTGA 2220  
Qy 2221 TATTTCATGCTGGATGTGTGGCATCCGGAACCTGACACCAAGAGAGAGAGAGAGAG 2280  
Db 2221 TATTTCATGCTGGATGTGTGGCATCCGGAACCTGACACCAAGAGAGAGAGAGAGAG 2280  
Qy 2281 CAATTTAGCATGAATTCATGCAAGCTTTGGGAACTCTGGAGAGA 2324  
Db 2281 CAATTTAGCATGAATTCATGCAAGCTTTGGGAACTCTGGAGAGA 2324

## RESULT 4

US-09-903-216-3  
; Sequence 3, Application US/09903216  
; Patent No. US2002011481A1  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV2

```
; CURRENT APPLICATION NUMBER: US/09/903,216
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/436,184
; PRIOR FILING DATE: 1999-11-08
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2324
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-903-216-3

Query Match      100.0%; Score 2324; DB 9; Length 2324;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGACCGTGCAATGGCCCGAGCGTAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCG 60
DB 1 CGGACCGTGCAATGGCCCGAGCGTAAGATGCCAAGAGCAGCGGCAACAGCAGCAGCG 60
QY 61 GCTCCGCGAGCGGTAGCAGCGTGGCGCAGCAGCGCCCGGGGCCCGGAGAGACAA 120
DB 61 GCTCCGCGAGCGGTAGCAGCGTGGCGCAGCAGCGCCCGGGGCCCGGAGAGACAA 120
QY 121 AGCATGGAGGACACAAGAAATGGGAGAAAGCGGACTCTCGGGAACTTTCATCTTCAGT 180
DB 121 AGCATGGAGGACACAAGAAATGGGAGAAAGCGGACTCTCGGGAACTTTCATCTTCAGT 180
QY 181 GGTTCATGGTGAATGCAATGCTGGCGCTGGCAACATGTAAGTGTGCTGTTGGTTGATC 240
DB 181 GGTTCATGGTGAATGCAATGCTGGCGCTGGCAACATGTAAGTGTGCTGTTGGTTGATC 240
QY 241 TTGTTGACTATGAGGAAGTCTAGGAAATAGGAAATAGGAAATAGGAAATAGGAAAT 300
DB 241 TTGTTGACTATGAGGAAGTCTAGGAAATAGGAAATAGGAAATAGGAAATAGGAAAT 300
QY 301 ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 301 ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 361 CAGCAGTCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 420
DB 361 CAGCAGTCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 420
QY 421 AGGCGAGAACCCAGAAATTCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCCATG 480
DB 421 AGGCGAGAACCCAGAAATTCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCCATG 480
QY 481 AAATGGTACACGCAAGCAGTGTGAGGAGAGAGACTTGCAACAGAGATGGACCCACAG 540
DB 481 AAATGGTACACGCAAGCAGTGTGAGGAGAGAGACTTGCAACAGAGATGGACCCACAG 540
QY 541 GAGAACCAACAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 541 GAGAACCAACAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 601 AGACCTCGAACTGAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGGAAGA 660
DB 601 AGACCTCGAACTGAAGTATCTCATGAAGAAACCGAGCATAGTTACCACTGGGAAGA 660
QY 661 CAGTTTCACAAGACTGTAATCAGGATATGGAAGATGGAAGATGGAAGATGGAAGATGGA 720
DB 661 CAGTTTCACAAGACTGTAATCAGGATATGGAAGATGGAAGATGGAAGATGGAAGATGGA 720
QY 721 ATTCAGTGAACAGTAGTAGAGATGAAGATTCACCATCATACAGATGATGATGATGAT 780
DB 721 ATTCAGTGAACAGTAGTAGAGATGAAGATTCACCATCATACAGATGATGATGATGATGAT 780
QY 781 ACCAAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAAAATGAAGGATAGAAATCA 840
DB 781 ACCAAGTCTATGAGGAACAGCAGTATATGAACCTCTAGAAAAATGAAGGATAGAAATCA 840
QY 841 CAGAAGTAACGTCTCCCTCGAGGATAATCTCTGTAGAAAGATTCACAGGTAATGTAGAG 900
```





1381 AAAATGACCTTGGCGTGGATACCTCTTGATAGAGATAAATGACAATCAAGAAGTTT 1440  
1381 AAAATGACCTTGGCGTGGATACCTCTTGATAGAGATAAATGACAATCAAGAAGTTT 1440  
1441 ATGAAGAGTCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTATGGCTTCA 1500  
1441 ATGAAGAGTCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTATGGCTTCA 1500  
1501 TCCTGAAGGCACAGAACAAAATTGCTGAGAGCATCCCATATTTAAAGGAAGAAATAGAAAT 1560  
1501 TCCTGAAGGCACAGAACAAAATTGCTGAGAGCATCCCATATTTAAAGGAAGAAATAGAAAT 1560  
1561 CCGGAGATCTCGGCACCTGATGAGAGATTTTATTTCCACTGGGGATGCCATGCGA 1620  
1561 CCGGAGATCTCGGCACCTGATGAGAGATTTTATTTCCACTGGGGATGCCATGCGA 1620  
1621 GGGTTGGGAAACAAGAGGCATATAAGTGGTATGAGCTTTGGGCACAAGAGAGGACACTTTG 1680  
1621 GGGTTGGGAAACAAGAGGCATATAAGTGGTATGAGCTTTGGGCACAAGAGAGGACACTTTG 1680  
1681 CATCTGCTGCGCAACCGCTCACTCTACAATGTGAATGGACTGAAGCAGACAGCTTTGGTGA 1740  
1681 CATCTGCTGCGCAACCGCTCACTCTACAATGTGAATGGACTGAAGCAGACAGCTTTGGTGA 1740  
1741 CCCCAAAAGAAACGGGCTTACACAGAGTTAGTAAAGTCTTTTAAAGAACTGGAAGTTAA 1800  
1741 CCCCAAAAGAAACGGGCTTACACAGAGTTAGTAAAGTCTTTTAAAGAACTGGAAGTTAA 1800  
1801 TCCGAGATGAAGGCTTTCAGTGTGATGATGAATGAAGCAAGGCTCTTCTCGCTCAGGATG 1860  
1801 TCCGAGATGAAGGCTTTCAGTGTGATGATGAATGAAGCAAGGCTCTTCTCGCTCAGGATG 1860  
1861 AAAACCTCAGGAAAAAGGGGACTGGAGCCAGTTTCACTGCTGGCAGCAAGAAAGAAAGAA 1920  
1861 AAAACCTCAGGAAAAAGGGGACTGGAGCCAGTTTCACTGCTGGCAGCAAGAAAGAAAGAA 1920  
1921 ATGAAATGCTGCAAGAGGCTCTTAAACCTGTACCTTAACTAGTAAAGTTTCCCGAGA 1980  
1921 ATGAAATGCTGCAAGAGGCTCTTAAACCTGTACCTTAACTAGTAAAGTTTCCCGAGA 1980  
1981 CAACAGGATGAGAGAGGACAGATCAAAATTTCCATCATGCACCCCGGACTCACGTTG 2040  
1981 CAACAGGATGAGAGAGGACAGATCAAAATTTCCATCATGCACCCCGGACTCACGTTG 2040  
2041 GCGCGCACACAGGGGCCCAAACTCGAGGCTCCGAATGCACTGCGCTTGGTGAATTCCTCA 2100  
2041 GCGCGCACACAGGGGCCCAAACTCGAGGCTCCGAATGCACTGCGCTTGGTGAATTCCTCA 2100  
2101 AGGAAGGCTGCAAGATTCGATGTGCGAACGAGACCCAGGACCTGGGAGGAAGGCTGC 2160  
2101 AGGAAGGCTGCAAGATTCGATGTGCGAACGAGACCCAGGACCTGGGAGGAAGGCTGC 2160  
2161 TCATCTTTGATGACTCTTTCGAGCAGAGGATGCGAGGATGCGCTCATCTTTCCGGCTGA 2220  
2161 TCATCTTTGATGACTCTTTCGAGCAGAGGATGCGAGGATGCGCTCATCTTTCCGGCTGA 2220  
2221 TATTTCATCGTGGATGTGGCATCCGGAATCGAACCTGACACACAGCAGAGACCGAGCTTCCAG 2280  
2221 TATTTCATCGTGGATGTGGCATCCGGAATCGGAATCGAACCTGACACACAGCAGAGACCGAGCTTCCAG 2280  
2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTCTGGAGAGA 2324  
2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAAACTCTGGAGAGA 2324

RESULT 6

US-09-903-023-3  
; Sequence 3, Application US/09903023  
; Patent No. US20020146421A1  
; GENERAL INFORMATION:  
; APPLICANT: Wanda, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim

; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV1  
; CURRENT APPLICATION NUMBER: US/09/903,023  
; CURRENT FILING DATE: 2001-10-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-903-023-3

Query Match 100.0%; Score 2324; DB 9; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCCGTGCAATGGCCCGCAGCGTAAGAAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
Db 1 CGGACCCGTGCAATGGCCCGCAGCGTAAGAAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60  
Qy 61 GCTCCGCGAGCGGTAGCACAGAGTCCGGGCGAGCAGCAGCGCCCGGGCCCGGAGAGAGACAA 120  
Db 61 GCTCCGCGAGCGGTAGCACAGAGTCCGGGCGAGCAGCAGCGCCCGGGCCCGGAGAGAGACAA 120  
Qy 121 AGCATGAGAGACACAAATGGGAGGAAAGCGGACTCTCGGGAACTTCTTCCTTTCACGT 180  
Db 121 AGCATGAGAGACACAAATGGGAGGAAAGCGGACTCTCGGGAACTTCTTCCTTTCACGT 180  
Qy 181 GGTTTATGGTGAATGCAATGCTGGGCGCTCTGGACATCTGTAGCTGTCTGTTGGTTGATC 240  
Db 181 GGTTTATGGTGAATGCAATGCTGGGCGCTCTGGACATCTGTAGCTGTCTGTTGGTTGATC 240  
Qy 241 TTGTTGACTATGAGGAAGTTCTAGGAAAATCTAGGAATCTATGATGCTGATGGTGGATGGAG 300  
Db 241 TTGTTGACTATGAGGAAGTTCTAGGAAAATCTAGGAATCTATGATGCTGATGGTGGATGGAG 300  
Qy 301 ATTTTGTATGGTGTGATGTCGCAAGTTTATTTAGGACTTAAAGAGAGATCTACTTTCAGAGC 360  
Db 301 ATTTTGTATGGTGTGATGTCGCAAGTTTATTTAGGACTTAAAGAGAGATCTACTTTCAGAGC 360  
Qy 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGACAGGTTCTCTGTGG 420  
Db 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGACAGGTTCTCTGTGG 420  
Qy 421 AGGAGAACCCCGAGAAATATCGAAGATGAAAGAAAGCAAAATTCAGTCCCTTCTCCATG 480  
Db 421 AGGAGAACCCCGAGAAATATCGAAGATGAAAGAAAGCAAAATTCAGTCCCTTCTCCATG 480  
Qy 481 AAATGGTACACGCGAGAAATCTTTGAGGAGAGAACTTGGCAAGAGAGATGGACCCACAG 540  
Db 481 AAATGGTACACGCGAGAAATCTTTGAGGAGAGAACTTGGCAAGAGAGATGGACCCACAG 540  
Qy 541 GAGAACCAACAGAGAGATGATGATTTCTTATGGGACTGATGATGATGATGATGATGATG 600  
Db 541 GAGAACCAACAGAGAGATGATGATTTCTTATGGGACTGATGATGATGATGATGATGATG 600  
Qy 601 AGACCCCTGGAACTGAAAGTATCTCATGAAGAAACCGAGCATAGTTTACCATGCGAAGAGA 660  
Db 601 AGACCCCTGGAACTGAAAGTATCTCATGAAGAAACCGAGCATAGTTTACCATGCGAAGAGA 660  
Qy 661 CAGTTTCAAGACTGTAAATCAGGATATGGAAGAGATGATGCTCTGAGCAGGAAATCCAG 720  
Db 661 CAGTTTCAAGACTGTAAATCAGGATATGGAAGAGATGATGCTCTGAGCAGGAAATCCAG 720  
Qy 721 ATTCCAGTGAACAGTATGAGAGATGAAGATTTGACCATGATGATGATGATGATGATGAT 780  
Db 721 ATTCCAGTGAACAGTATGAGAGATGAAGATTTGACCATGATGATGATGATGATGATGAT 780  
Qy 781 ACCAAGTCTATGAGGACACAGCAGTATATGAACCTCTAGAAAATGAAGGGATAGAAATCA 840  
Db 781 ACCAAGTCTATGAGGACACAGCAGTATATGAACCTCTAGAAAATGAAGGGATAGAAATCA 840



Db 781 ACCAAGTCTATGAGAAACAAGCAGTATATGAACCTCTAGAAAATGAAGGATAGAATCA 840  
Qy 841 CAGAGTAATCTGCCCTCCTGAGGATAATCTCTGTAGAGATTCAAGGTAAATTGTGAAG 900  
Db 841 CAGAGTAATCTGCCCTCCTGAGGATAATCTCTGTAGAGATTCAAGGTAAATTGTGAAG 900  
Qy 901 AAGTAAGCATTTTCTGTGTGAAGAAACAGCAGGAGAGTACCACCGAAGCAAAATAGAAAAA 960  
Db 901 AAGTAAGCATTTTCTGTGTGAAGAAACAGCAGGAGAGTACCACCGAAGCAAAATAGAAAAA 960  
Qy 961 CAGATGATCCAGAACAAAAAGCAAAAGTTAAAGAAAAGAGCCTAAATCTTTAAATAAT 1020  
Db 961 CAGATGATCCAGAACAAAAAGCAAAAGTTAAAGAAAAGAGCCTAAATCTTTAAATAAT 1020  
Qy 1021 TTGATAAGCACTATTAAAGCTGAATCTTGAATCTGCGAAGAAATCTCGTAAAGGGAAAAA 1080  
Db 1021 TTGATAAGCACTATTAAAGCTGAATCTTGAATCTGCGAAGAAATCTCGTAAAGGGAAAAA 1080  
Qy 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGCAATAACCTCAGAGTCCACGAG 1140  
Db 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGCAATAACCTCAGAGTCCACGAG 1140  
Qy 1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAAATGAGG 1200  
Db 1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAAGAGGAGAAATGAGG 1200  
Qy 1201 TGCTACGTGGAGCCATCGACACCTACCAAGAGTGGCCAGCTACCTGATGTCCTCGCAG 1260  
Db 1201 TGCTACGTGGAGCCATCGACACCTACCAAGAGTGGCCAGCTACCTGATGTCCTCGCAG 1260  
Qy 1261 ACCTGCTGAAGCTGATTTGAAGCTGCTCAGACAGGCAAAATTTCTAGGTCATATGA 1320  
Db 1261 ACCTGCTGAAGCTGATTTGAAGCTGCTCAGACAGGCAAAATTTCTAGGTCATATGA 1320  
Qy 1321 GAGGTTCCCTCTTACCCTCAGAGATTAGTTCAACTATTTCCCAATGATACTTCTCTAA 1380  
Db 1321 GAGGTTCCCTCTTACCCTCAGAGATTAGTTCAACTATTTCCCAATGATACTTCTCTAA 1380  
Qy 1381 AAAATGACCTTTGGCGTGGGATACCTCTGATAGGAGATAATGACAATGCAAGAAAGATTT 1440  
Db 1381 AAAATGACCTTTGGCGTGGGATACCTCTGATAGGAGATAATGACAATGCAAGAAAGATTT 1440  
Qy 1441 ATGAGAGTCTCAGTGTGACACCTAATCATGCGCTTTGCTAAAGCTCAATATGCGTTCA 1500  
Db 1441 ATGAGAGTCTCAGTGTGACACCTAATCATGCGCTTTGCTAAAGCTCAATATGCGTTCA 1500  
Qy 1501 TCCTGAAGGCAAGAAACAAATTTGCTGAGAGCATCCCATATTTAAAGAGGAAGTAAT 1560  
Db 1501 TCCTGAAGGCAAGAAACAAATTTGCTGAGAGCATCCCATATTTAAAGAGGAAGTAAT 1560  
Qy 1561 CCGGAGATCTCGGCACTGATGGGAGATTTTATTTCCACCTGGGGGATGCCATGAGA 1620  
Db 1561 CCGGAGATCTCGGCACTGATGGGAGATTTTATTTCCACCTGGGGGATGCCATGAGA 1620  
Qy 1621 GGGTTGGCAACAAAGAGGCATATAAGTGTATGAGCTTGGGCAACAGAGAGGACCTTTG 1680  
Db 1621 GGGTTGGCAACAAAGAGGCATATAAGTGTATGAGCTTGGGCAACAGAGAGGACCTTTG 1680  
Qy 1681 CATCTGCTGCAACGCTCACTCTACAATGTGAATGGACTGAAGAGCACGCTTGTGGA 1740  
Db 1681 CATCTGCTGCAACGCTCACTCTACAATGTGAATGGACTGAAGAGCACGCTTGTGGA 1740  
Qy 1741 CCCCCAAGAAACGGGCTACACAGAGTTAGTAAGTCTTTTAAAGAAAACTGGAAGTTAA 1800  
Db 1741 CCCCCAAGAAACGGGCTACACAGAGTTAGTAAGTCTTTTAAAGAAAACTGGAAGTTAA 1800  
Qy 1801 TCCGAGATGAAGGCTTTCAGATGATGATAAAGCAAAAGGCTCTTCTCCTCGTGGAGTG 1860  
Db 1801 TCCGAGATGAAGGCTTTCAGATGATGATAAAGCAAAAGGCTCTTCTCCTCGTGGAGTG 1860  
Qy 1861 AAAACCTGAGGAAGGAGGAGTGGAGCCAGTTCAAGCTGTGGCAGCAGGAGAGAA 1920  
Db 1861 AAAACCTGAGGAAGGAGGAGTGGAGCCAGTTCAAGCTGTGGCAGCAGGAGAGAA 1920

Qy 1921 ATGAAATGCTCTCAAGAGAGCTCTTAAAACTGTACCTTACTAGAAAAAGTTTCCCGAGA 1980  
Db 1921 ATGAAATGCTCTCAAGAGAGCTCTTAAAACTGTACCTTACTAGAAAAAGTTTCCCGAGA 1980  
Qy 1981 CAACAGGATGAGAGAGAGACAGATCAAAATATTTCATCATGACACCCCGGAGCTACGTTG 2040  
Db 1981 CAACAGGATGAGAGAGAGACAGATCAAAATATTTCATCATGACACCCCGGAGCTACGTTG 2040  
Qy 2041 GSCCGCACACAGGCGCCACAACTGAGGCTCCGAATGCACTGGGCTTGTGATTTCCCA 2100  
Db 2041 GSCCGCACACAGGCGCCACAACTGAGGCTCCGAATGCACTGGGCTTGTGATTTCCCA 2100  
Qy 2101 AGGAGGCTGCAAGATTCGATGTGCCAGAGACAGGACCTGGGAGGAGGCAAGTGC 2160  
Db 2101 AGGAGGCTGCAAGATTCGATGTGCCAGAGACAGGACCTGGGAGGAGGCAAGTGC 2160  
Qy 2161 TCATCTTTGATGACTCTTTTGAACAGAGGATGAGGATGCTCTCATCTTTTCCGGCTGA 2220  
Db 2161 TCATCTTTGATGACTCTTTTGAACAGAGGATGAGGATGCTCTCATCTTTTCCGGCTGA 2220  
Qy 2221 TATTTCATCGTGGATGTGTGGCATCCGGAATCTGACACCAACAGCAGAGAGCTTCCAG 2280  
Db 2221 TATTTCATCGTGGATGTGTGGCATCCGGAATCTGACACCAACAGCAGAGAGCTTCCAG 2280  
Qy 2281 CAATTTAGCTGAATTCATCAAGCTTGGGAACTCTCGAGAGA 2324  
Db 2281 CAATTTAGCTGAATTCATCAAGCTTGGGAACTCTCGAGAGA 2324

## RESULT 7

US-09-436-184-3  
; Sequence 3, Application US/09436184  
; Publication No. US20030031670A1  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: R. I. Hoep. - Malignant Neoplasms  
; CURRENT APPLICATION NUMBER: US/09/436,184  
; CURRENT FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-436-184-3

Query Match 100.0%; Score 2324; DB 10; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCGTGCATGGCCCGCGTAAGATGCCAAGCAGCGCGCAACAGCAGCAGCGCG 60  
Db 1 CGGACCGTGCATGGCCCGCGTAAGATGCCAAGCAGCGCGCGCAACAGCAGCAGCGCG 60  
Qy 61 GCTCCGGCAGCGGTAGCAGCGAGTGGCGGAGCAGAGCCCGGGGCGCGGAGAGACAA 120  
Db 61 GCTCCGGCAGCGGTAGCAGCGAGTGGCGGAGCAGAGCCCGGGGCGCGGAGAGACAA 120  
Qy 121 AGCATGGAGGACACAAGATGGGAGGAAAGGCGGACTCTCGGGAACCTTCACTTCA 180  
Db 121 AGCATGGAGGACACAAGATGGGAGGAAAGGCGGACTCTCGGGAACCTTCACTTCA 180  
Qy 181 GGTATTGATGATGATGCTGGCGCTGCGACATCTGTAGCTGTCTGTTGTTGATC 240  
Db 181 GGTATTGATGATGATGCTGGCGCTGCGACATCTGTAGCTGTCTGTTGTTGATC 240  
Qy 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGCTGCTGCTGCTGCTG 300  
Db 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGCTGCTGCTGCTGCTG 300

Db 241 TTGTTGACTATGAGGAAGTTCTTAGGAAAACTAGGAATCTATGATGCTGATGGTATGGAG 300  
Qy 301 ATTTTGTATGTGATGATGCCAAAGTTTATTATAGACTTTAAAGAGAGATCTACTTTCAGAGC 360  
Db 301 ATTTTGTATGTGATGATGCCAAAGTTTATTATAGACTTTAAAGAGAGATCTACTTTCAGAGC 360  
Qy 361 CAGCAGTCCCAGGAGAGGCTGAGCCACACACTGAGCCCCGAGGAGCAGTTTCTCTGTGG 420  
Db 361 CAGCAGTCCCAGGAGAGGCTGAGCCACACACTGAGCCCCGAGGAGCAGTTTCTCTGTGG 420  
Qy 421 AGGAGAAACCCAGAGATTCGAAGATGAAGCAAAAGAACTTCAATTCAGTCCCTTCTCCATG 480  
Db 421 AGGAGAAACCCAGAGATTCGAAGATGAAGCAAAAGAACTTCAATTCAGTCCCTTCTCCATG 480  
Qy 481 AAATGGTACGCGAGAACATGTTGAGGAGAGAGCTTGCACAAAGAGATGGACCCACAG 540  
Db 481 AAATGGTACGCGAGAACATGTTGAGGAGAGAGCTTGCACAAAGAGATGGACCCACAG 540  
Qy 541 GAGAACCAACAAAGAGATGATGAGTTTCTTATGGCGACTGATGTAGATGATGATTTG 600  
Db 541 GAGAACCAACAAAGAGATGATGAGTTTCTTATGGCGACTGATGTAGATGATGATTTG 600  
Qy 601 AGACCTCGAACTCTGAAGATCTCATGAAGAAACCGAGCATAGTTACACGTGGAAGAGA 660  
Db 601 AGACCTCGAACTCTGAAGATCTCATGAAGAAACCGAGCATAGTTACACGTGGAAGAGA 660  
Qy 661 CAGTTTCAAGACTGTAAATCAGGATATGGAAGAGATGATCTGAGCAGGAAATCCAG 720  
Db 661 CAGTTTCAAGACTGTAAATCAGGATATGGAAGAGATGATCTGAGCAGGAAATCCAG 720  
Qy 721 ATTCCAGTGAACCAAGATAGAAAGATGACCATGATACAGATGATGTAACAT 780  
Db 721 ATTCCAGTGAACCAAGATAGAAAGATGACCATGATACAGATGATGTAACAT 780  
Qy 781 ACCAAGTCTATGAGAAACAGCAGTATATGAACCTCTAGAAATGAAGGGATAGAAATCA 840  
Db 781 ACCAAGTCTATGAGAAACAGCAGTATATGAACCTCTAGAAATGAAGGGATAGAAATCA 840  
Qy 841 CAGAGTAATCTCCCTCGAGGATAATCTGTAGAGATTCACAGTAAATTTAGAG 900  
Db 841 CAGAGTAATCTCCCTCGAGGATAATCTGTAGAGATTCACAGTAAATTTAGAG 900  
Qy 901 AAGTAAGCATTTTCTCTGGAAGAACAGCAGGAGTACCACAGAAACAAATAGAAAA 960  
Db 901 AAGTAAGCATTTTCTCTGGAAGAACAGCAGGAGTACCACAGAAACAAATAGAAAA 960  
Qy 961 CAGATGATCCAGAACAAAGCAAAAGTTAGAAAGAACGCTTAACTTTAAATAAT 1020  
Db 961 CAGATGATCCAGAACAAAGCAAAAGTTAGAAAGAACGCTTAACTTTAAATAAT 1020  
Qy 1021 TTGATAAGCACTATTAAAGCTGAACCTTGATGCTGCAGAAAACTCCGTAAAAAGGGAAAA 1080  
Db 1021 TTGATAAGCACTATTAAAGCTGAACCTTGATGCTGCAGAAAACTCCGTAAAAAGGGAAAA 1080  
Qy 1081 TTGAGGAAGCAGTGAATGATTTAAAGAACTAGTAGCAAAATACCCCTCAGAGTCCACGAG 1140  
Db 1081 TTGAGGAAGCAGTGAATGATTTAAAGAACTAGTAGCAAAATACCCCTCAGAGTCCACGAG 1140  
Qy 1141 CAAGATATGGAGGCGCAGTGTAGGATGATTTGGCTGAGAGAGGAGAGTAAATGAGG 1200  
Db 1141 CAAGATATGGAGGCGCAGTGTAGGATGATTTGGCTGAGAGAGGAGAGTAAATGAGG 1200  
Qy 1201 TGCTACGTGGAGCCATCCAGACCTACCAGAGGTGGCCAGCCTACTCATGTCCCTGCAG 1260  
Db 1201 TGCTACGTGGAGCCATCCAGACCTACCAGAGGTGGCCAGCCTACTCATGTCCCTGCAG 1260  
Qy 1261 ACCTGCTGAAGCTGAGTTGAAGCGCTCGCTCAGACAGGCAACAAATTTCTAGTCAATGA 1320  
Db 1261 ACCTGCTGAAGCTGAGTTGAAGCGCTCGCTCAGACAGGCAACAAATTTCTAGTCAATGA 1320  
Qy 1321 GAGGTTCCCTGCTTACCCCTGAGAGATGATTTCAACTATTTCCCAATGATACCTCTTAA 1380  
Db 1321 GAGGTTCCCTGCTTACCCCTGAGAGATGATTTCAACTATTTCCCAATGATACCTCTTAA 1380

Qy 1381 AAAATGACCTTGGCGTGGGATACCTCTTGTAGAGAGATAATGACAAATGCAAGAAAGTTT 1440  
Db 1381 AAAATGACCTTGGCGTGGGATACCTCTTGTAGAGAGATAATGACAAATGCAAGAAAGTTT 1440  
Qy 1441 ATGAAGAGGTGCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATATATGGCTTCA 1500  
Db 1441 ATGAAGAGGTGCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATATATGGCTTCA 1500  
Qy 1501 TCCTGAAGGACACAGAACAAATTTGCTGAGAGCATCCCATATTTTAAAGGAAGGAATAGAAT 1560  
Db 1501 TCCTGAAGGACACAGAACAAATTTGCTGAGAGCATCCCATATTTTAAAGGAAGGAATAGAAT 1560  
Qy 1561 CCGGAGATCTCTGGCAGCTGATGAGGGAGATTTTATTTCCACCTGGGGGATGCCATGCGA 1620  
Db 1561 CCGGAGATCTCTGGCAGCTGATGAGGGAGATTTTATTTCCACCTGGGGGATGCCATGCGA 1620  
Qy 1621 GGGTTGGGAACAAAGAGGCTATATAGTGTGATGAGCTTTGGGCACAAAGAGAGGACACTTTG 1680  
Db 1621 GGGTTGGGAACAAAGAGGCTATATAGTGTGATGAGCTTTGGGCACAAAGAGAGGACACTTTG 1680  
Qy 1681 CATCTGTCTGCAACGCTCACTCTACAAATGTGAATGACCTGAAAGCAAGCCTTGGTGA 1740  
Db 1681 CATCTGTCTGCAACGCTCACTCTACAAATGTGAATGACCTGAAAGCAAGCCTTGGTGA 1740  
Qy 1741 CCCCAGGAGAAACGGGCTCACACAGAGTTAGTAAAGTCTTTAGAAAGAACTTGGAAAGTAA 1800  
Db 1741 CCCCAGGAGAAACGGGCTCACACAGAGTTAGTAAAGTCTTTAGAAAGAACTTGGAAAGTAA 1800  
Qy 1801 TCCGAGATGAAGGCTTTCAGTGTGATGAATGAAGCAAGGCTCTCTTCTGCTGAGGATG 1860  
Db 1801 TCCGAGATGAAGGCTTTCAGTGTGATGAATGAAGCAAGGCTCTCTTCTGCTGAGGATG 1860  
Qy 1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTTCAGCTGTGCGAGCAAGGAAAGAGAA 1920  
Db 1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTTCAGCTGTGCGAGCAAGGAAAGAGAA 1920  
Qy 1921 ATGAAAAATGCTGCAAAAGGAGCTCTTAAACCTGTACCTTACTAGAAAAAGTTCCCGGAGA 1980  
Db 1921 ATGAAAAATGCTGCAAAAGGAGCTCTTAAACCTGTACCTTACTAGAAAAAGTTCCCGGAGA 1980  
Qy 1981 CAAAGAGATGAGAGAGGAGAGATCAAAATATTCATCATGCAACCCCGGAGTCACTGT 2040  
Db 1981 CAAAGAGATGAGAGAGGAGAGATCAAAATATTCATCATGCAACCCCGGAGTCACTGT 2040  
Qy 2041 GGCCGCAACAGGCCCCCAAACTGCAAGCTCCGAGTCCGAAATGCACTGGGCTTGGTGAATCCCA 2100  
Db 2041 GGCCGCAACAGGCCCCCAAACTGCAAGCTCCGAGTCCGAAATGCACTGGGCTTGGTGAATCCCA 2100  
Qy 2101 AGGAAGGCTGCAAGATTCGATGTGCCAAACGAGACCCAGGACCTGGAGAGGAGGCAAGGTGC 2160  
Db 2101 AGGAAGGCTGCAAGATTCGATGTGCCAAACGAGACCCAGGACCTGGAGAGGAGGCAAGGTGC 2160  
Qy 2161 TCATCTTTGATGACTCTTTTGGAGCAGGATGATGGAGGATGCCTCATCTTTCGGGCTGA 2220  
Db 2161 TCATCTTTGATGACTCTTTTGGAGCAGGATGATGGAGGATGCCTCATCTTTCGGGCTGA 2220  
Qy 2221 TATTTCATGCTGGAGTGTGGCATCCGAACTGACCAACAGCAGAGAGAGGAGGCTTCCAG 2280  
Db 2221 TATTTCATGCTGGAGTGTGGCATCCGAACTGACCAACAGCAGAGAGAGGAGGCTTCCAG 2280  
Qy 2281 CAATTTAGCATGAATTCATGCAAGCTTTGGGAAACTCTGGAGAGA 2324  
Db 2281 CAATTTAGCATGAATTCATGCAAGCTTTGGGAAACTCTGGAGAGA 2324

## RESULT 8

US-10-084-817-63  
; Sequence 63, Application US/10084817  
; Publication No. US20030119009A1  
; GENERAL INFORMATION:  
; APPLICANT: Susan Stuart  
; APPLICANT: Jed G. Nuchtern

```

; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shonett
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 63
; LENGTH: 5358
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 089172.13
US-10-084-817-63

Query Match      93.4%; Score 2171; DB 15; Length 5358;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2321; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1  CGGACCGTGCATGGCCAGCGTAAGAAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60
DB      258 CGGACCGTGCATGGCCAGCGTAAGAAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 317

QY      61  GCTCCGGCAGCGGTAGCAGCAGTCCGGCAGCAGCAGCCCGGGCCCGGAGAGACAA 120
DB      318 GCTCCGGCAGCGGTAGCAGCAGTCCGGCAGCAGCAGCCCGGGCCCGGAGAGACAA 377

QY      121 AGCATGGAGGACACAGAATGGAGGAAAGCGGACTCTCGGGAACTTCAATCTTCACGT 180
DB      378 AGCATGGAGGACACAGAATGGAGGAAAGCGGACTCTCGGGAACTTCAATCTTCACGT 437

QY      181 GGTATTATGGTATTCATTCGCGGCGTCTGGACATCTGTAGCTGTGCTGTTGGTTGATC 240
DB      438 GGTATTATGGTATTCATTCGCGGCGTCTGGACATCTGTAGCTGTGCTGTTGGTTGATC 497

QY      241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGGTATGGAG 300
DB      498 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGATGGTATGGAG 557

QY      301 ATTGTGATGGATGATGCCAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 360
DB      558 ATTGTGATGGATGATGCCAAGTTTATTAGGACTTAAAGAGAGATCTACTTCAGAGC 617

QY      361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTGTGG 420
DB      618 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCGAGGAGCAGGTTCTGTGG 677

QY      421 AGGCAGAACCCCGAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480
DB      678 AGGCAGAACCCCGAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 737

QY      481 AAATGGTACACGACAGACATGTTGAGGAGAGAGACTTGCACAGAGAGATGGACCCACAG 540
DB      738 AAATGGTACACGACAGACATGTTGAGGAGAGAGACTTGCACAGAGAGATGGACCCACAG 797

QY      541 GAGAACCAACAAAGAGGATGATGAGTTTCTTATGGCACTGATGTAGATGATGATTTG 600
DB      798 GAGAACCAACAAAGAGGATGATGAGTTTCTTATGGCACTGATGTAGATGATGATTTG 857

QY      601 AGACCTGGAACTGAAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTCGGAAGAGA 660
DB      858 AGACCTGGAACTGAAAGTATCTCATGAAGAAACCGAGCATAGTTACACGTCGGAAGAGA 917

QY      661 CAGTTTCACAGACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720
DB      918 CAGTTTCACAGACTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 977

QY      721 ATTCCAGTGAACAGTAGTAGAGATGAAGATTCACCATGATACAGATGATGATTAACAT 780
DB

```

```

DB      978 ATTCAGTGAACAGTAGTAGAGATGAAGAGATTGCACCATGATACAGATGATGTAACAT 1037
QY      781 ACCAAGTCTATGAGGAAACAAGCAGTATATGAAACCTCTAGAAAATGAAGGATGAGAAATCA 840
DB      1038 ACCAAGTCTATGAGGAAACAAGCAGTATATGAAACCTCTAGAAAATGAAGGATGAGAAATCA 1097
QY      841 CAGAAGTAACCTGCTCCCGCTGAGGATAATCTCTGTAGAAGATTACAGAGTAATTTGTAGAAG 900
DB      1098 CAGAAGTAACCTGCTCCCGCTGAGGATAATCTCTGTAGAAGATTACAGAGTAATTTGTAGAAG 1157
QY      901 AAGTAAGCATTTTCTCTGTGGAAGAACACAGCAGAGTAGTACCACAGAAACAAATAGAAAAA 960
DB      1158 AAGTAAGCATTTTCTCTGTGGAAGAACACAGCAGAGTAGTACCACAGAAACAAATAGAAAAA 1217
QY      961 CAGATGATCCAGAACAAAAGGTAAGAAAAGAGAGCGCTTAAACTTTTAAATAAT 1020
DB      1218 CAGATGATCCAGAACAAAAGGTAAGAAAAGAGAGCGCTTAAACTTTTAAATAAT 1277
QY      1021 TTGATAAGACTATTAAAGCTGAACTTTGATGCTGCAAAAAAACTCCGTAAGAAGGGGAAAAA 1080
DB      1278 TTGATAAGACTATTAAAGCTGAACTTTGATGCTGCAAAAAAACTCCGTAAGAAGGGGAAAAA 1337
QY      1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGCGCAAAATACCTCAGAGTCCACGAG 1140
DB      1338 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGCGCAAAATACCTCAGAGTCCACGAG 1397
QY      1141 CAAGATATGGGAAGGCGCAGTGTGAGGATGATTTGCGCTGAGAGAGAGAGTAATGAGG 1200
DB      1398 CAAGATATGGGAAGGCGCAGTGTGAGGATGATTTGCGCTGAGAGAGAGAGTAATGAGG 1457
QY      1201 TGCTAGTGGAGCCATCGAGACCTTCAAGAGGTTGGCCAGCTTACCTGATGTCCTGCGAG 1260
DB      1458 TGCTAGTGGAGCCATCGAGACCTTCAAGAGGTTGGCCAGCTTACCTGATGTCCTGCGAG 1517
QY      1261 ACTGCTGAAGCTGAGTTTGAAGCGTGTGCTCAGACAGAGCAACAAATTTCTAGGTCTATGA 1320
DB      1518 ACTGCTGAAGCTGAGTTTGAAGCGTGTGCTCAGACAGAGCAACAAATTTCTAGGTCTATGA 1577
QY      1321 GAGGTTCCCTGCTTACCCTGCGAGATTAGTTTCAACTATTTCCCAATGATACTTCCCTTAA 1380
DB      1578 GAGGTTCCCTGCTTACCCTGCGAGATTAGTTTCAACTATTTCCCAATGATACTTCCCTTAA 1637
QY      1381 AAAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATAATGCAATGCAAGAAAGATTT 1440
DB      1638 AAAATGACCTTGGCGTGGGATACCTCTTGATAGGAGATAATGCAATGCAAGAAAGATTT 1697
QY      1441 ATGAAGAGGTTGCTGAGTGTGACACCTTAATGATGGCTTTGCTTAAAGTCCATTTATGGCTTCA 1500
DB      1698 ATGAAGAGGTTGCTGAGTGTGACACCTTAATGATGGCTTTGCTTAAAGTCCATTTATGGCTTCA 1757
QY      1501 TCCTGAAGGCGACAGAACAAATTTGCTGAGAGCATCCCATATTTAAAGAAAGGAATAGAAT 1560
DB      1758 TCCTGAAGGCGACAGAACAAATTTGCTGAGAGCATCCCATATTTAAAGAAAGGAATAGAAT 1817
QY      1561 CGGAGATCTTGGCAGCTGATGAGGAGATTTTATTTCCCTGCGGATGCGCATGAGAG 1620
DB      1818 CGGAGATCTTGGCAGCTGATGAGGAGATTTTATTTCCCTGCGGATGCGCATGAGAG 1877
QY      1621 GGGTTGGGACAAAGAGGCAATATAGTGTGTAGCTTTGGGCAACAAGAGGACACTTTG 1680
DB      1878 GGGTTGGGACAAAGAGGCAATATAGTGTGTAGCTTTGGGCAACAAGAGGACACTTTG 1937
QY      1681 CATCTGCTGGCAACCGCTCACTCTCAATGTGAATGGACTGAAAGCAGACCCCTTGGTGA 1740
DB      1938 CATCTGCTGGCAACCGCTCACTCTCAATGTGAATGGACTGAAAGCAGACCCCTTGGTGA 1997
QY      1741 CCCAAAAGAAAAACGGGCTACACAGAGTGTAGTAAAGTCTTTTAAAGAAAGAACTGGAAGTTAA 1800
DB      1998 CCCAAAAGAAAAACGGGCTACACAGAGTGTAGTAAAGTCTTTTAAAGAAAGAACTGGAAGTTAA 2057
QY      1801 TCCGAGATGAGGCGCTTGGAGTGTAGTAAAGCAAAAGGCTCTCTTCCCTGCGAGGATG 1860
DB      2058 TCCGAGATGAGGCGCTTGGAGTGTAGTAAAGCAAAAGGCTCTCTTCCCTGCGAGGATG 2117

```

Qy 1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTTCAAGCTGTGGCAGCAAGGAAGAAGAA 1920  
Db 2118 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTTCAAGCTGTGGCAGCAAGGAAGAAGAA 2177  
Qy 1921 ATGAAATGCTGCAAGAGGAGCTCTAAACCTGTACCTTACTAGAAAAGTTCCCGGAGA 1980  
Db 2178 ATGAAATGCTGCAAGAGGAGCTCTAAACCTGTACCTTACTAGAAAAGTTCCCGGAGA 2237  
Qy 1981 CAACAGGATGAGAGAGGACAGATCAAAATATTCATCATGACACCCGGGACTCACGTGT 2040  
Db 2238 CAACAGGATGAGAGAGGACAGATCAAAATATTCATCATGACACCCGGGACTCACGTGT 2297  
Qy 2041 GGCCGACACAGGGGCCCAAACTGCAAGCTCCGAATGACCTGGCTTGTGTATCCCA 2100  
Db 2298 GGCCGACACAGGGGCCCAAACTGCAAGCTCCGAATGACCTGGCTTGTGTATCCCA 2357  
Qy 2101 AGGAAGGCTGCAAGATTCATGTGCTCAACAGGACCCAGGACTGGGAGGAGGCAAGGTGC 2160  
Db 2358 AGGAAGGCTGCAAGATTCATGTGCTCAACAGGACCCAGGACTGGGAGGAGGCAAGGTGC 2417  
Qy 2161 TCATCTTTGATGACTCCTTTGAGCAGCAGGTATGGCAGGATGCCTCATCTTTCCGGCTGA 2220  
Db 2418 TCATCTTTGATGACTCCTTTGAGCAGCAGGTATGGCAGGATGCCTCATCTTTCCGGCTGA 2477  
Qy 2221 TATTCATCGTGGATGTGTGGCATCCGGAATCGAACCTGGAACCTTGGAGAGA 2324  
Db 2478 TATTCATCGTGGATGTGTGGCATCCGGAATCGAACCTGGAACCTTGGAGAGA 2591  
RESULT 9  
US-10-096-534-37  
; Sequence 37, Application US/10096534  
; Publication No. US20030166887A1  
; GENERAL INFORMATION:  
; APPLICANT: The Brigham and Women's Hospital, Inc.  
; APPLICANT: Yates, Karen  
; APPLICANT: Mizuno, Shuichi  
; APPLICANT: Glowacki, Julie  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF SKELETAL DEGENERATION CONDITIONS  
; FILE REFERENCE: B0801/7244/KA/ERP  
; CURRENT APPLICATION NUMBER: US/10/096.534  
; CURRENT FILING DATE: 2002-03-12  
; PRIOR APPLICATION NUMBER: US 60/274,980  
; PRIOR FILING DATE: 2001-03-12  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 37  
; LENGTH: 2680  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-096-534-37  
Query Match 38.5%; Score 895; DB 16; Length 2680;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 945; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 CGGACCGTGCAATGGCCAGCGTAAGAATGCCAAGAGCAGCGGCAACAGCAGCAGCG 60  
Db 1 CGGACCGTGCAATGGCCAGCGTAAGAATGCCAAGAGCAGCGGCAACAGCAGCAGCG 60  
Qy 61 GCTCCGCGAGCGGTAGCAGTGGCGCAGCAGCAGCCCGGGCCCGGAGAGGACAA 120  
Db 61 GCTCCGCGAGCGGTAGCAGTGGCGCAGCAGCAGCCCGGGCCCGGAGAGGACAA 120  
Qy 121 AGCATGGAGGACACAAGAATGGAGAGGAGCGGACTCTCGGGAACTTCACTTTACGT 180  
Db 121 AGCATGGAGGACACAAGAATGGAGAGGAGCGGACTCTCGGGAACTTCACTTTACGT 180

Qy 181 GGTTTATGGTGATTCATTCCTGGCGCTCTGGACATCTGTAGCTGTCTGTTGGTTGATC 240  
Db 181 GGTTTATGGTGATTCATTCCTGGCGCTCTGGACATCTGTAGCTGTCTGTTGGTTGATC 240  
Qy 241 TTGTTGACTATGAGGAAGTTCTAGAAAACTAGAAATCTATGATGCTGCTGATGGTATGAG 300  
Db 241 TTGTTGACTATGAGGAAGTTCTAGAAAACTAGAAATCTATGATGCTGCTGATGGTATGAG 300  
Qy 301 ATTTTGTATGGTATGCTCCAAAGTTTTTATTAGGACTTAAAGAGAGATCTACTTCCAG 360  
Db 301 ATTTTGTATGGTATGCTCCAAAGTTTTTATTAGGACTTAAAGAGAGATCTACTTCCAG 360  
Qy 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGACAGGTTCTCTGG 420  
Db 361 CAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGGAGGACAGGTTCTCTGG 420  
Qy 421 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
Db 421 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
Qy 481 AAATGGTACACGCGAAGCATGTTTGGGAGAGAATCTTGCAACAAGAGATGGACCCACAG 540  
Db 481 AAATGGTACACGCGAAGCATGTTTGGGAGAGAATCTTGCAACAAGAGATGGACCCACAG 540  
Qy 541 GAGAACCAACAACAGAGATGATGATGTTCTTATGCGCATGATGTAGTATGATGATTTG 600  
Db 541 GAGAACCAACAACAGAGATGATGATGTTCTTATGCGCATGATGTAGTATGATGATTTG 600  
Qy 601 AGACCTTGGAACTGAAATGATCTCATGAAGAAACCGAGCATAGTTACACGCTGGAAGAGA 660  
Db 601 AGACCTTGGAACTGAAATGATCTCATGAAGAAACCGAGCATAGTTACACGCTGGAAGAGA 660  
Qy 661 CAGTTTCAAGACTGTAATCAGGATATGGAAGAGATGATGTTCTGAGCAGGAAATCCAG 720  
Db 661 CAGTTTCAAGACTGTAATCAGGATATGGAAGAGATGATGTTCTGAGCAGGAAATCCAG 720  
Qy 721 ATTCCAGTGAACAGTAGTAGAAGATGAAGATGACCATGATGATACAGATGATGTAACAT 780  
Db 721 ATTCCAGTGAACAGTAGTAGAAGATGAAGATGACCATGATGATACAGATGATGTAACAT 780  
Qy 781 ACCAAGTCTATGAGGAACAACAGCATATATGAACCTCTAGAAAAATGAAGGGATAGAAATCA 840  
Db 781 ACCAAGTCTATGAGGAACAACAGCATATATGAACCTCTAGAAAAATGAAGGGATAGAAATCA 840  
Qy 841 CAGAAGTAACCTGCTCCCGCTGAGGATAATCTCTGTAAGAGATTTCAGGTAATTTGTAAG 900  
Db 841 CAGAAGTAACCTGCTCCCGCTGAGGATAATCTCTGTAAGAGATTTCAGGTAATTTGTAAG 900  
Qy 901 AAGTAAGCATTTTCTCTGTGGAAGAACAGCAGGAGGATGATGATGATGATGATGATGAT 946  
Db 901 AAGTAAGCATTTTCTCTGTGGAAGAACAGCAGGAGGATGATGATGATGATGATGATGAT 946  
RESULT 10  
US-10-269-909-3  
; Sequence 3, Application US/10269909  
; Publication No. US20030180747A1  
; GENERAL INFORMATION:  
; APPLICANT: HRUBAN, RALPH H.  
; APPLICANT: ARGANI, PEDRAM  
; APPLICANT: IACOBIZIO-DONAHUE, CHRISTINE  
; APPLICANT: MAITRA, ANIRBAN  
; TITLE OF INVENTION: PANCREATIC CANCER DIAGNOSIS AND THERAPIES  
; FILE REFERENCE: 58303(71699)  
; CURRENT APPLICATION NUMBER: US/10/269,909  
; CURRENT FILING DATE: 2003-10-11  
; PRIOR APPLICATION NUMBER: 60/328,609  
; PRIOR FILING DATE: 2001-10-11  
; PRIOR APPLICATION NUMBER: 60/332,754  
; PRIOR FILING DATE: 2001-11-19  
; NUMBER OF SEQ ID NOS: 87  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3







Db 406 CAAGAAGATGGACCCACAGGAGAACACAAAGAGGATGATGAGTTCTTATGGCGACT 465  
Qy 582 GATGTAGATGATAGATTGAGACCTCGAACCTGGAATATCTCATGAGAAACCGAGCAT 641  
Db 466 GATGTAGATGATAGATTGAGACCTCGAACCTGGAATATCTCATGAGAAACCGAGCAT 525  
Qy 642 AGTTACCAAGTGGAGAGACAGTTTCACAGACTGTATCAGGATATGGAAGATGATG 701  
Db 526 AGTTACCAAGTGGAGAGACAGTTTCACAGACTGTATCAGGATATGGAAGATGATG 585  
Qy 702 TCTGAGCAGGAAATTCAGATTCAGTGAACACAGTAGTAGAAGATGAAAGATTGCACCAT 761  
Db 586 TCTGAGCAGGAAATTCAGATTCAGTGAACACAGTAGTAGAAGATGAAAGATTGCACCAT 645  
Qy 762 GATACAGATGATGTAACATACCAAGTCTATGAGAACCAAGCAGTAGTAGAAGATTGCACCAT 821  
Db 646 GATACAGATGATGTAACATACCAAGTCTATGAGAACCAAGCAGTAGTAGAAGATTGCACCAT 705  
Qy 822 AATGAAGGATAGAAATTCACAGAAGTAACCTG 852  
Db 706 AATGAAGGATAGAAATTCACAGAAGTAACCTG 736

## RESULT 14

US-10-017-754-1791  
; Sequence 1791, Application US/10017754  
; Publication No. US20030054363A1  
; GENERAL INFORMATION:  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Wang, Tonglong  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Johnson, Jeffrey C.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Warnerakis, Margarita  
; APPLICANT: Carter, Derrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: McNabb, Andria  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.478C18  
; CURRENT APPLICATION NUMBER: US/10/017,754  
; NUMBER OF SEQ ID NOS: 2004  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1791  
; LENGTH: 2442  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-017-754-1791

Query Match 27.5%; Score 640; DB 14; Length 2442;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 162 GGAATCTCATCTTTCACGTGGTTTATGGTGATTCATGCTGGCGCTCGACATCTGTA 221  
Db 46 GGAATCTCATCTTTCACGTGGTTTATGGTGATTCATGCTGGCGCTCGACATCTGTA 105  
Qy 222 GCTGTGTTGGTTTGTATCTTGTGACTATGAGGAAGTCTAGGAAACTAGGAATCTAT 281  
Db 106 GCTGTGTTGGTTTGTATCTTGTGACTATGAGGAAGTCTAGGAAACTAGGAATCTAT 165  
Qy 282 GATGCTGATGATGAGATTTTGTATGATGATGATGATGATGATGATGATGATGATGAT 341  
Db 166 GATGCTGATGATGAGATTTTGTATGATGATGATGATGATGATGATGATGATGATGAT 225  
Qy 342 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCC 401  
Db 226 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCCAGAGAGGCTGAGCCACACACTGAGCCC 285  
Qy 402 GAGGAGCAGGTTCTGTGGAGGAGAGACCCAGATATTCAGAGATGAGCAAGAGACAA 461

Db 286 GAGGAGCAGGTTCTGTGGAGGAGAGACCCAGAAATATCGAAGATGAGCAAAAGACAA 345  
Qy 462 ATTCACTCCCTTCTCCATGAATGGTACACGAGAACATGTTGAGGAGAGAGACTTGCAG 521  
Db 346 ATTCACTCCCTTCTCCATGAATGGTACACGAGAACATGTTGAGGAGAGAGACTTGCAG 405  
Qy 522 CAAGAAGATGAGACCCACAGAGAACCAACCAAGAGGATGATGAGTTCTTATGGCGACT 581  
Db 406 CAAGAAGATGAGACCCACAGAGAACCAACCAAGAGGATGATGAGTTCTTATGGCGACT 465  
Qy 582 GATGTAGATGATAGATTGAGACCTCGAACCTGGAATATCTCATGAGAAACCGAGCAT 641  
Db 466 GATGTAGATGATAGATTGAGACCTCGAACCTGGAATATCTCATGAGAAACCGAGCAT 525  
Qy 642 AGTTACCAAGTGGAGAGACAGTTTCACAGACTGTATCAGGATATGGAAGATGATG 701  
Db 526 AGTTACCAAGTGGAGAGACAGTTTCACAGACTGTATCAGGATATGGAAGATGATG 585  
Qy 702 TCTGAGCAGGAAATTCAGATTCAGTGAACCAAGTAGTAGAAGATGAAAGATTGCACCAT 761  
Db 586 TCTGAGCAGGAAATTCAGATTCAGTGAACCAAGTAGTAGAAGATGAAAGATTGCACCAT 645  
Qy 762 GATACAGATGATGTAACATACCAAGTCTATGAGAACCAAGCAGTAGTAGAAGATTGCACCAT 821  
Db 646 GATACAGATGATGTAACATACCAAGTCTATGAGAACCAAGCAGTAGTAGAAGATTGCACCAT 705  
Qy 822 AATGAAGGATAGAAATTCACAGAAGTAACCTG 852  
Db 706 AATGAAGGATAGAAATTCACAGAAGTAACCTG 736

## RESULT 15

US-10-113-872-1791  
; Sequence 1791, Application US/10113872  
; Publication No. US20030170255A1  
; GENERAL INFORMATION:  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Sleath, Paul R.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Derrick  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.478C19  
; CURRENT APPLICATION NUMBER: US/10/113,872  
; NUMBER OF SEQ ID NOS: 2011  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1791  
; LENGTH: 2442  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-113-872-1791

Query Match 27.5%; Score 640; DB 16; Length 2442;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 162 GGAATCTCATCTTTCACGTGGTTTATGGTGATTCATGCTGGCGCTCGACATCTGTA 221  
Db 46 GGAATCTCATCTTTCACGTGGTTTATGGTGATTCATGCTGGCGCTCGACATCTGTA 105  
Qy 222 GCTGTGTTGGTTTGTATCTTGTGACTATGAGGAAGTCTAGGAAACTAGGAATCTAT 281  
Db 106 GCTGTGTTGGTTTGTATCTTGTGACTATGAGGAAGTCTAGGAAACTAGGAATCTAT 165  
Qy 282 GATGCTGATGATGAGATTTTGTATGATGATGATGATGATGATGATGATGATGATGAT 341  
Db 166 GATGCTGATGATGAGATTTTGTATGATGATGATGATGATGATGATGATGATGATGAT 225







```

; LENGTH: 660
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 507..554, 599, 656
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-3202

Query Match      15.4%; Score 359; DB 11; Length 660;
Best Local Similarity 100.0%; Pred. No. 1.1e-174;
Matches 359; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 581 TGATGTAGATAGATTTGAGACCTCGAACCTCGAAGTATCTCATGAAGAAACCGAGCA 640
Db 462 TCATGTAGATAGATTTGAGACCTCGAACCTCGAAGTATCTCATGAAGAAACCGAGCA 403

QY 641 TAGTTACCACGTGGAAGAGACAGTTTCAAGAGACTGTAATCAGGATATGGAAGAGATGAT 700
Db 402 TAGTTACCACGTGGAAGAGACAGTTTCAAGAGACTGTAATCAGGATATGGAAGAGATGAT 343

QY 701 GTCTGAGCAGGAAATCCAGATTCAGTGAAACCGAGTAGTAGAAGATGAAGATTCACCA 760
Db 342 GTCTGAGCAGGAAATCCAGATTCAGTGAAACCGAGTAGTAGAAGATGAAGATTCACCA 283

QY 761 TGATACAGATGATGTAACATACCACTCTATGAGGAACAGCAGTATATGAACCTCTAGA 820
Db 282 TGATACAGATGATGTAACATACCACTCTATGAGGAACAGCAGTATATGAACCTCTAGA 223

QY 821 AAATGAAGGATAGAAATACACAGAGTAACCTGCTCCCTCGAGGATATCTCTGAGAAGA 880
Db 222 AAATGAAGGATAGAAATACACAGAGTAACCTGCTCCCTCGAGGATATCTCTGAGAAGA 163

QY 881 TTCACAGTAATGTAGAGAAGTAAGCATTTTCTCTGGAAGAACAGCAGGAGATAC 939
Db 162 TTCACAGTAATGTAGAGAAGTAAGCATTTTCTCTGGAAGAACAGCAGGAGATAC 104

RESULT 20
US-10-242-535A-25661
; Sequence 25661, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR FILING DATE: 2001-07-13
; PRIOR FILING DATE: 2001-03-12
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25661
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-25661

Query Match      13.6%; Score 315; DB 17; Length 366;
Best Local Similarity 99.7%; Pred. No. 6.4e-152;
Matches 365; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 191 GATTGCATTGCTGGCGCTCTGGACATCTGTAGCTGTCGTTTGGTTTGTGATCTTTGACTA 250
Db 1 GATTGCATTGCTGGCGCTCTGGACATCTGTAGCTGTCGTTTGGTTTGTGATCTTTGACTA 60

QY 251 TGAGGAAGTTCTAGGAAACCTAGCAATCTATGATGCTGATGGTGTAGGATTTGATGT 310
Db 61 TGAGGAAGTTCTAGGAAACCTAGCAATCTATGATGCTGATGGTGTAGGATTTGATGT 120

QY 311 GGATGATGCCAAAGTTTATTATTAGGACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCC 370
Db 121 GGATGATGCCAAAGTTTATTATTAGGACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCC 180

QY 371 GCCAGAGAGGCTGAGCCACACACTTGAGCCCGAGGAGCAGGTTCCTGTGGAGGAGCAAC 430
Db 371 GCCAGAGAGGCTGAGCCACACACTTGAGCCCGAGGAGCAGGTTCCTGTGGAGGAGCAAC 430

QY 431 CCAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATGAATGGTACA 490
Db 241 CCAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATGAATGGTACA 300

Query Match      13.6%; Score 315; DB 17; Length 366;
Best Local Similarity 99.7%; Pred. No. 6.4e-152;
Matches 365; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 191 GATTGCATTGCTGGCGCTCTGGACATCTGTAGCTGTCGTTTGGTTTGTGATCTTTGACTA 250
Db 1 GATTGCATTGCTGGCGCTCTGGACATCTGTAGCTGTCGTTTGGTTTGTGATCTTTGACTA 60

QY 251 TGAGGAAGTTCTAGGAAACCTAGCAATCTATGATGCTGATGGTGTAGGATTTGATGT 310
Db 61 TGAGGAAGTTCTAGGAAACCTAGCAATCTATGATGCTGATGGTGTAGGATTTGATGT 120

QY 311 GGATGATGCCAAAGTTTATTATTAGGACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCC 370
Db 121 GGATGATGCCAAAGTTTATTATTAGGACTTAAAGAGAGATCTACTTCAGAGCCAGCAGTCCC 180

QY 371 GCCAGAGAGGCTGAGCCACACACTTGAGCCCGAGGAGCAGGTTCCTGTGGAGGAGCAAC 430
Db 371 GCCAGAGAGGCTGAGCCACACACTTGAGCCCGAGGAGCAGGTTCCTGTGGAGGAGCAAC 430

QY 431 CCAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATGAATGGTACA 490
Db 241 CCAGAAATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATGAATGGTACA 300

```

QY 491 CGCAGAACATGTTGAGGAGAGACTTGCACAGAGATGGACCCACAGGAGAACCA 550  
| | | | |  
Db 301 CGCAGAACATGTTGAGGAGAGACTTGCACAGAGATGGACCCACAGGAGAACCA 360  
| | | | |  
QY 551 ACAGA 556  
| | | | |  
Db 361 ACAGA 366  
| | | | |

## RESULT 22

US-10-066-543-277  
; Sequence 277, Application US/10066543  
; Publication No. US20030087818A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yuqiu  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Darriek  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066.543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 277  
; LENGTH: 495  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 341, 343  
; OTHER INFORMATION: n = A,T,C or G  
US-10-066-543-277

Query Match 10.7%; Score 249; DB 14; Length 495;  
Best Local Similarity 99.7%; Pred. No. 1e-117;  
Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 630 GAAACCGAGCATGTTACACGCTGGAAGAGACAGTTTCACAGACTGTTATCAGGATG 689  
| | | | |  
Db 41 GAAACCGAGCATGTTACACGCTGGAAGAGACAGTTTCACAGACTGTTATCAGGATG 100  
| | | | |  
QY 690 GAAGAGATGATCTGAGCAGGAAATCCAGATTCAGTGAAACAGTAGTAGAAGTAA 749  
| | | | |  
Db 101 GAAGAGATGATCTGAGCAGGAAATCCAGATTCAGTGAAACAGTAGTAGAAGTAA 160  
| | | | |  
QY 750 AGATTGCCACATGATACAGATGATGTACATACCAAGTCTATGAGGAACAAGCAGTATAT 809  
| | | | |  
Db 161 AGATTGCCACATGATACAGATGATGTACATACCAAGTCTATGAGGAACAAGCAGTATAT 220  
| | | | |  
QY 810 GAACTCTAGAAATGAAGGATAGAAATACAGAGTAACCTGCTCCCTGAGGATAT 869  
| | | | |  
Db 221 GAACTCTAGAAATGAAGGATAGAAATACAGAGTAACCTGCTCCCTGAGGATAT 280  
| | | | |  
QY 870 CTTGTAGAGATTTCACAGGTAATTTGAGAGAGTAAGCATTTTCTGTGGAAGACAG 929  
| | | | |  
Db 281 CTTGTAGAGATTTCACAGGTAATTTGAGAGAGTAAGCATTTTCTGTGGAAGACAG 340  
| | | | |

## RESULT 23

US-10-242-535A-2707  
; Sequence 2707, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:

APPLICANT: ChondroGene Inc.  
APPLICANT: Liew, C.C.  
TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
FILE REFERENCE: 4231/2005  
CURRENT APPLICATION NUMBER: US/10/242.535A  
CURRENT FILING DATE: 2002-09-12  
PRIOR APPLICATION NUMBER: US 10/085,783  
PRIOR FILING DATE: 2002-02-28  
PRIOR APPLICATION NUMBER: US 60/305,340  
PRIOR FILING DATE: 2001-07-13  
PRIOR APPLICATION NUMBER: US 60/275,017  
PRIOR FILING DATE: 2001-03-12  
PRIOR APPLICATION NUMBER: US 60/271,955  
PRIOR FILING DATE: 2001-02-28  
NUMBER OF SEQ ID NOS: 58994  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 2707  
LENGTH: 191  
TYPE: DNA  
ORGANISM: Human  
US-10-242-535A-2707

Query Match 5.8%; Score 135; DB 17; Length 191;  
Best Local Similarity 100.0%; Pred. No. 1.1e-58;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 367 TCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGGAGGCG 426  
| | | | |  
Db 1 TCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGGAGGCG 60  
| | | | |  
QY 427 AACCCAGATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATGAATGG 486  
| | | | |  
Db 61 AACCCAGATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATGAATGG 120  
| | | | |  
QY 487 TACACGCGAGAACATG 501  
| | | | |  
Db 121 TACACGCGAGAACATG 135  
| | | | |

## RESULT 24

US-10-085-783A-2707  
; Sequence 2707, Application US/10085783A  
; Publication No. US20040037841A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2707  
; LENGTH: 191  
; TYPE: DNA  
; ORGANISM: Human  
US-10-085-783A-2707

Query Match 5.8%; Score 135; DB 17; Length 191;  
Best Local Similarity 100.0%; Pred. No. 1.1e-58;  
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 367 TCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGGAGGCG 426  
| | | | |  
Db 1 TCCGCCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGGTTCTCTGGAGGCG 60  
| | | | |  
QY 427 AACCCAGATATCGAAGATGAAGCAAAAGAACAAATTCAGTCCCTTCTCCATGAATGG 486  
| | | | |



```
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 2825
; LENGTH: 450
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-2825

Query Match      3.4%; Score 78; DB 17; Length 450;
Best Local Similarity 100.0%; Pred. No. 3.7e-29;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 349 CTACTTCAGAGCCAGCAGTCCGCCAGAGAGGCTGAGCCACACTGAGCCGAGGAGC 408
DB 450 CTACTTCAGAGCCAGCAGTCCGCCAGAGAGGCTGAGCCACACTGAGCCGAGGAGC 391

QY 409 AGGTTCTCTGGAGGCGAG 426
DB 390 AGGTTCTCTGGAGGCGAG 373

RESULT 29
US-10-443-622-24
; Sequence 24, Application US/10443622
; Publication No. US20040024192A1
; GENERAL INFORMATION:
; APPLICANT: Carter et al.
; TITLE OF INVENTION: 19 Human Secreted Proteins
; FILE REFERENCE: PZ009P1
; CURRENT APPLICATION NUMBER: US/10/443,622
; CURRENT FILING DATE: 2003-05-23
; PRIOR APPLICATION NUMBER: PCT/US98/13608
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: 60/051,480
; PRIOR FILING DATE: 1997-07-01
; PRIOR APPLICATION NUMBER: 60/051,381
; PRIOR FILING DATE: 1997-07-01
; PRIOR APPLICATION NUMBER: 60/058,663
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,598
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 24
; LENGTH: 1382
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-443-622-24

Query Match      3.1%; Score 72; DB 17; Length 1382;
Best Local Similarity 100.0%; Pred. No. 5e-26;
Matches 72; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 263 AGGAAACATAGGAATCTATGATGCTGATGCGAGATTTTGGATGGATGATGCCAA 322
DB 233 AGGAAACATAGGAATCTATGATGCTGATGCGAGATTTTGGATGGATGATGCCAA 292

QY 323 AGTTTATTAGG 334
```

```
DB 293 AGTTTATTAGG 304

RESULT 30
US-09-908-975-11070
; Sequence 11070, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 11070
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-11070

Query Match      2.6%; Score 60; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 7.3e-20;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 672 GACTGTATCAGGATATGGAAGAGATGATGCTGAGCAGGAGAAATCCAGTGAA 731
DB 1 GACTGTATCAGGATATGGAAGAGATGATGCTGAGCAGGAGAAATCCAGTGAA 60

RESULT 31
US-10-116-275-322/c
; Sequence 322, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 322
; LENGTH: 2208
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-116-275-322

Query Match      2.5%; Score 59; DB 17; Length 2208;
Best Local Similarity 100.0%; Pred. No. 2.8e-19;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2239 GGCATCCGGAACTGACACACAGAGAGCCAGCCTTCCAGCAATTTAGCATGATTC 2297
DB 59 GGCATCCGGAACTGACACACAGAGAGCCAGCCTTCCAGCAATTTAGCATGATTC 1
```





## US-10-027-632-7020

Query Match 1.7%; Score 44; DB 17; Length 591;  
Best Local Similarity 100.0%; Pred. No. 1.5e-11; Indels 0; Gaps 0;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 800 ACCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCACAG 843  
Db 304 ACCAGTATATGAACCTCTAGAAAATGAAGGATAGAAATCACAG 347

## RESULT 36

US-10-027-632-56205/c  
; Sequence 56205, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720-  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 56205  
; LENGTH: 453  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-56205

Query Match 1.7%; Score 39; DB 13; Length 453;  
Best Local Similarity 100.0%; Pred. No. 6e-09;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 349 CTACTTCAGCCAGCAGTCCCGCCAGAGAGGCTGAGC 387  
Db 453 CTACTTCAGCCAGCAGTCCCGCCAGAGAGGCTGAGC 415

## RESULT 37

US-10-027-632-310460/c  
; Sequence 310460, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 310460  
; LENGTH: 453  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-310460

Query Match 1.7%; Score 39; DB 13; Length 453;  
Best Local Similarity 100.0%; Pred. No. 6e-09;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 349 CTACTTCAGCCAGCAGTCCCGCCAGAGAGGCTGAGC 387  
Db 453 CTACTTCAGCCAGCAGTCCCGCCAGAGAGGCTGAGC 415

## RESULT 38

US-10-027-632-56205/c  
; Sequence 56205, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 56205  
; LENGTH: 453  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-56205

Query Match 1.7%; Score 39; DB 17; Length 453;  
Best Local Similarity 100.0%; Pred. No. 6e-09;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 349 CTACTTCAGCCAGCAGTCCCGCCAGAGAGGCTGAGC 387  
Db 453 CTACTTCAGCCAGCAGTCCCGCCAGAGAGGCTGAGC 415

## RESULT 39

US-10-027-632-310460/c  
; Sequence 310460, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 310460  
; LENGTH: 453  
; TYPE: DNA  
; ORGANISM: Human  
; US-10-632-310460

Query Match 1.7%; Score 39; DB 17; Length 453;  
Best Local Similarity 100.0%; Pred. No. 6e-09;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 349 CTACTTCAGCCAGCAGTCCCGCCAGAGAGGCTGAGC 387  
Db 453 CTACTTCAGCCAGCAGTCCCGCCAGAGAGGCTGAGC 415

## RESULT 40

US-10-424-599-97119  
; Sequence 97119, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 97119  
; LENGTH: 859  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_58710C.1  
US-10-424-599-97119

Query Match 1.0%; Score 24; DB 17; Length 859;  
Best Local Similarity 100.0%; Pred. No. 0.36;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 AGCAGCGCAACAGCAGCAGCAGC 59  
Db 128 AGCAGCGCAACAGCAGCAGCAGC 151

## RESULT 41

US-10-425-115-124491/c  
; Sequence 124491, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 124491  
; LENGTH: 2135  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_45013C.1  
US-10-425-115-124491

Query Match 1.0%; Score 23; DB 18; Length 2135;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 37 GCAGCGGCAACAGCAGCAGCAGC 59  
Db 1991 GCAGCGGCAACAGCAGCAGCAGC 1969

## RESULT 42

US-10-437-963-73067  
; Sequence 73067, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 73067  
; LENGTH: 681  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_73384C.1  
US-10-437-963-73067

Query Match 0.9%; Score 21; DB 18; Length 681;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 CGGCAACAGCAGCAGCAGCAGC 61  
Db 39 CGGCAACAGCAGCAGCAGCAGC 59

## RESULT 43

US-10-425-115-42949/c  
; Sequence 42949, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115

```
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 363326
; SEQ ID NO 42949
; LENGTH: 715
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577.139178C.1
US-10-425-115-42949

Query Match      0.9%; Score 21; DB 18; Length 715;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 GCGGCAACAGCAGCAGCG 60
Db 462 GCGGCAACAGCAGCAGCG 442

RESULT 44
US-10-308-503-188
; Sequence 188, Application US/10308503
; Publication No. US20030191080A1
; GENERAL INFORMATION:
; APPLICANT: PHILLIPS, M. IAN
; TITLE OF INVENTION: ANTISENSE COMPOSITIONS TARGETED TO BETAL-ADRENOCEPTOR-SPECIFIC mRNA
; FILE REFERENCE: 4300.013900
; CURRENT APPLICATION NUMBER: US/10/308,503
; PRIOR FILING DATE: 2003-12-03
; PRIOR APPLICATION NUMBER: US/09/614,034
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 09/152,717
; PRIOR FILING DATE: 1998-09-14
; PRIOR APPLICATION NUMBER: PCT/US99/21007
; PRIOR FILING DATE: 1999-09-14
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 188
; TYPE: DNA
; LENGTH: 1845
; ORGANISM: UNKNOWN
; FEATURE:
; OTHER INFORMATION: SYNTHETIC OLIGONUCLEOTIDE
; NAME/KEY: misc feature
; LOCATION: (1828)..(1828)
; OTHER INFORMATION: WHERE N = A, T, C OR G
US-10-308-503-188

Query Match      0.9%; Score 21; DB 16; Length 1845;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 CAGCAGCGGCTCCGCGCAGCG 73
Db 93 CAGCAGCGGCTCCGCGCAGCG 113

RESULT 45
US-09-859-604-10/c
; Sequence 10, Application US/09859604
; Patent No. US20020110559A1
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R.
; APPLICANT: de la Monte, Suzanne M
; APPLICANT: Ghanbari, Hossein A
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS
; FILE REFERENCE: 21486-032 CIP
; CURRENT APPLICATION NUMBER: US/09/859,604
; PRIOR FILING DATE: 2001-05-17
; OTHER INFORMATION:
; NUMBER OF SEQ ID NOS: 13
```

```
; PRIOR APPLICATION NUMBER: 09/436,184
; PRIOR FILING DATE: 1999-11-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Location (-1)
; OTHER INFORMATION: oligonucleotide
US-09-859-604-10

Query Match      0.9%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 AATGCCCCAGCGTAAGAATG 30
Db 20 AATGCCCCAGCGTAAGAATG 1

RESULT 46
US-09-859-604-11/c
; Sequence 11, Application US/09859604
; Patent No. US20020110559A1
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R.
; APPLICANT: de la Monte, Suzanne M
; APPLICANT: Deutch, Alan H
; APPLICANT: Ghanbari, Hossein A
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS
; FILE REFERENCE: 21486-032 CIP
; CURRENT APPLICATION NUMBER: US/09/859,604
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/436,184
; PRIOR FILING DATE: 1999-11-08
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; TYPE: DNA
; LENGTH: 20
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Location (-6)
; OTHER INFORMATION: oligonucleotide
US-09-859-604-11

Query Match      0.9%; Score 20; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CGTCAATGGCCAGCGTAA 25
Db 20 CGTCAATGGCCAGCGTAA 1

RESULT 47
US-09-859-604-12/c
; Sequence 12, Application US/09859604
; Patent No. US20020110559A1
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R.
; APPLICANT: de la Monte, Suzanne M
; APPLICANT: Deutch, Alan H
; APPLICANT: Ghanbari, Hossein A
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS
; FILE REFERENCE: 21486-032 CIP
; CURRENT APPLICATION NUMBER: US/09/859,604
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/436,184
; PRIOR FILING DATE: 1999-11-08
; NUMBER OF SEQ ID NOS: 13
```

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 12  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:Locations (-11)  
; OTHER INFORMATION: oligonucleotide  
US-09-859-604-12

Query Match 0.9%; Score 20; DB 9; Length 20;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCGTGCATGGCCAG 20  
Db 20 CGGACCGTGCATGGCCAG 1

RESULT 48

US-10-425-115-21348/c  
; Sequence 21348, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 21348  
; LENGTH: 183  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MFT4577\_119464C.1  
US-10-425-115-21348

Query Match 0.9%; Score 20; DB 18; Length 183;  
Best Local Similarity 100.0%; Pred. No. 40;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 79 CGAGTGGCGGCGAGCAGC 98  
Db 59 CGAGTGGCGGCGAGCAGC 40

RESULT 49

US-10-425-115-22449  
; Sequence 22449, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 22449  
; LENGTH: 432  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MFT4577\_120475C.1  
US-10-425-115-22449

Query Match 0.9%; Score 20; DB 18; Length 432;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 GCGGCAACAGCAGCAGC 59  
Db 396 GCGGCAACAGCAGCAGC 415

RESULT 50

US-10-767-701-26002/c  
; Sequence 26002, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10/767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 26002  
; LENGTH: 489  
; TYPE: DNA  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 30969203  
US-10-767-701-26002

Query Match 0.9%; Score 20; DB 18; Length 489;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 AGCAGCAGCGGCTCCGCGAG 70  
Db 242 AGCAGCAGCGGCTCCGCGAG 223

RESULT 51

US-10-424-599-89196  
; Sequence 89196, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 89196  
; LENGTH: 495  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(495)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_51552C.1  
US-10-424-599-89196

Query Match 0.9%; Score 20; DB 17; Length 495;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 991 AGAAAAAGAGCCCTAACTT 1010



```
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 322023
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-322023

Query Match          0.9%; Score 20; DB 13; Length 567;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 972 GAACAAAAGCAAAAGTTAA 991
Db 212 GAACAAAAGCAAAAGTTAA 193

RESULT 56
US-10-027-632-82527/c
; Sequence 82527, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 82527
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-82527

Query Match          0.9%; Score 20; DB 17; Length 567;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 972 GAACAAAAGCAAAAGTTAA 991
Db 212 GAACAAAAGCAAAAGTTAA 193

RESULT 57
US-10-027-632-322022/c
; Sequence 322022, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```

```
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 322022
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-322022

Query Match          0.9%; Score 20; DB 17; Length 567;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 972 GAACAAAAGCAAAAGTTAA 991
Db 212 GAACAAAAGCAAAAGTTAA 193

RESULT 58
US-10-027-632-322023/c
; Sequence 322023, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 322023
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-322023

Query Match          0.9%; Score 20; DB 17; Length 567;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 972 GAACAAAAGCAAAAGTTAA 991
Db 212 GAACAAAAGCAAAAGTTAA 193
```

Db 212 GAACAAAGCAAAAGTTAA 193

## RESULT 59

US-10-425-115-112661/c  
; Sequence 112661, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: LI, GANG  
; TITLE OF INVENTION: RAAV VECTOR-BASED PRO-OPIOMELANOCORTIN COMPOSITIONS AND METHODS  
; FILE REFERENCE: 4300.015400  
; CURRENT APPLICATION NUMBER: US/10/822,613  
; CURRENT FILING DATE: 2004-04-12  
; PRIOR FILING DATE: 2003-04-11  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 804  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-822-613-1  
; OTHER INFORMATION: Clone ID: MRT4577\_34237C.1

Query Match 0.9%; Score 20; DB 18; Length 750;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1933 GCAAGGAGCTCCTAAACC 1952  
Db 412 GCAAGGAGCTCCTAAACC 393

## RESULT 60

US-10-437-963-71261  
; Sequence 71261, Application US/10437963  
; Publication No. US2004012343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 71261  
; LENGTH: 771  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_71751C.1  
US-10-437-963-71261

Query Match 0.9%; Score 20; DB 18; Length 771;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCGGC 62  
Db 372 GCAACAGCAGCAGCGGC 391

## RESULT 61

US-10-822-613-1

; Sequence 1, Application US/10822613  
; Publication No. US20050002905A1  
; GENERAL INFORMATION:  
; APPLICANT: SCARPACE, PHILIP J.  
; APPLICANT: LI, GANG  
; TITLE OF INVENTION: RAAV VECTOR-BASED PRO-OPIOMELANOCORTIN COMPOSITIONS AND METHODS  
; FILE REFERENCE: 4300.015400  
; CURRENT APPLICATION NUMBER: US/10/822,613  
; CURRENT FILING DATE: 2004-04-12  
; PRIOR FILING DATE: 2003-04-11  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 804  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-822-613-1

Query Match 0.9%; Score 20; DB 18; Length 804;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCGGC 62  
Db 269 GCAACAGCAGCAGCGGC 288

## RESULT 62

US-10-437-963-24406  
; Sequence 24406, Application US/10437963  
; Publication No. US2004012343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 24406  
; LENGTH: 843  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_29392C.1  
US-10-437-963-24406

Query Match 0.9%; Score 20; DB 18; Length 843;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 54 AGCAGCGCTCCGCGCGG 73  
Db 366 AGCAGCGCTCCGCGCGG 385

## RESULT 63

US-10-437-963-51503  
; Sequence 51503, Application US/10437963  
; Publication No. US2004012343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua



```
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 51503
; LENGTH: 924
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_53889C.1
US-10-437-963-51503

Query Match          0.9%; Score 20; DB 18; Length 924;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 GCGCAACAGCAGCAGCAGC 59
Db 663 GCGCAACAGCAGCAGCAGC 682

RESULT 64
US-10-646-620A-8
; Sequence 8, Application US/10646620A
; Publication No. US20040138115A1
; GENERAL INFORMATION:
; APPLICANT: WITTIG, Burghardt
; APPLICANT: STEIN, Christoph
; APPLICANT: SCHAEFER, Michael
; APPLICANT: SCHROFF, Matthias
; APPLICANT: JUNGHANS, Claas
; APPLICANT: KOENIG MEREDIZ, Sven A.
; TITLE OF INVENTION: Local Pain Combating-Agent
; FILE REFERENCE: NHL-NP-43
; CURRENT APPLICATION NUMBER: US/10/646,620A
; CURRENT FILING DATE: 2003-08-22
; PRIOR FILING DATE: 2001-09-09
; PRIOR FILING DATE: 2001-02-24
; PRIOR FILING DATE: 2002-02-19
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 936
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Intron
; LOCATION: (847)..(933)
; OTHER INFORMATION: beta-endorphin cDNA sequence
US-10-646-620A-8

Query Match          0.9%; Score 20; DB 18; Length 936;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCAGCGGC 62
Db 401 GCAACAGCAGCAGCAGCGGC 420

RESULT 65
US-10-425-115-180542/c
; Sequence 180542, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc. et al.
; TITLE OF INVENTION: 95 Human secreted proteins
; FILE REFERENCE: PZ027P1
; CURRENT APPLICATION NUMBER: US/09/397,945
```

```
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 180542
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1011)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_96234C.1
US-10-425-115-180542

Query Match          0.9%; Score 20; DB 18; Length 1011;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 AGCAGCAGCGCTCCGGCAG 70
Db 229 AGCAGCAGCGCTCCGGCAG 210

RESULT 66
US-09-853-386-105
; Sequence 105, Application US/09853386
; Patent No. US20020049151A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Evelyn
; APPLICANT: Bresnihan, Barry
; APPLICANT: Conneely, Orla
; APPLICANT: Fitzgerald, Oliver
; TITLE OF INVENTION: Therapeutic Approaches to Diseases by Suppression of the NURR
; FILE REFERENCE: P01972US1
; CURRENT APPLICATION NUMBER: US/09/853,386
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/203645
; PRIOR FILING DATE: 2000-05-12
; NUMBER OF SEQ ID NOS: 153
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 105
; LENGTH: 1072
; TYPE: DNA
; ORGANISM: HUMAN
; OTHER INFORMATION:
US-09-853-386-105

Query Match          0.9%; Score 20; DB 9; Length 1072;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCAGCGGC 62
Db 375 GCAACAGCAGCAGCAGCGGC 394

RESULT 67
US-09-397-945-80/c
; Sequence 80, Application US/09397945
; Publication No. US20030065139A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc. et al.
; TITLE OF INVENTION: 95 Human secreted proteins
; FILE REFERENCE: PZ027P1
; CURRENT APPLICATION NUMBER: US/09/397,945
```

```
; CURRENT FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: PCI/US99/05804
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/078,566
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,576
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,573
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,574
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,579
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/080,314
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080,312
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/078,578
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,581
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,577
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,563
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/080,313
; PRIOR FILING DATE: 1998-04-01
; NUMBER OF SEQ ID NOS: 470
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 80
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-397-945-80

Query Match          0.9%; Score 20; DB 10; Length 1109;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      43  GCAACAGCAGCAGCAGCGGC 62
Db      152  GCACAGCAGCAGCAGCGGC 133

RESULT 68
US-10-653-595-80/c
; Sequence 80, Application US/10653595
; Publication No. US20040048304A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et. al.
; TITLE OF INVENTION: 98 Human secreted proteins
; FILE REFERENCE: P2027P1C8
; CURRENT APPLICATION NUMBER: US/10/653,595
; PRIOR FILING DATE: 2003-09-03/397945
; PRIOR APPLICATION NUMBER: US 03/397945
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: PCI/US99/05804
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/078,566
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,576
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,573
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,574
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,579
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/080,314
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080,312
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/078,578
```

```
; PRIOR FILING DATE: 1998-03-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 80
; LENGTH: 1109
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-653-595-80

Query Match          0.9%; Score 20; DB 17; Length 1109;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      43  GCAACAGCAGCAGCAGCGGC 62
Db      152  GCAACAGCAGCAGCAGCGGC 133

RESULT 69
US-09-960-288-1
; Sequence 1, Application US/09960288
; Patent No. US20020064833A1
; GENERAL INFORMATION:
; APPLICANT: Fagerlund, Tore
; APPLICANT: Alestrom, Peter
; APPLICANT: Berg, K+re
; TITLE OF INVENTION: Recombinant plasmids and method for treating substance abuse
; FILE REFERENCE: 406801
; CURRENT APPLICATION NUMBER: US/09/960,288
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/426,877
; PRIOR FILING DATE: 1999-10-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-960-288-1

Query Match          0.9%; Score 20; DB 9; Length 1230;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      43  GCAACAGCAGCAGCAGCGGC 62
Db      474  GCAACAGCAGCAGCAGCGGC 493

RESULT 70
US-10-287-971-37/c
; Sequence 37, Application US/10287971
; Publication No. US20040067882A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook, et al
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-480A
; CURRENT APPLICATION NUMBER: US/10/287,971
; CURRENT FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: 09/997,425
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: 10/035,568
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 60/338,626
; PRIOR FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: 60/401,479
; PRIOR FILING DATE: 2002-08-06
; PRIOR APPLICATION NUMBER: 60/333,072
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 60/348,283
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: 60/393,262
```

;  
; PRIOR FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: 60/406,181  
; PRIOR FILING DATE: 2002-08-26  
; NUMBER OF SEQ ID NOS: 397  
; SOFTWARE: CuraSeqList version 0.1  
; SEQ ID NO 37  
; LENGTH: 1245  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (90)..(1091)  
US-10-287-971-37

Query Match 0.9%; Score 20; DB 17; Length 1245;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCAGCGGC 62  
Db 138 GCAACAGCAGCAGCAGCGGC 119

RESULT 71  
US-10-425-114-23644  
; Sequence 23644, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 23644  
; LENGTH: 1282  
; TYPE: DNA  
; ORGANISM: Oryza sativa nipponbare  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3599-031-D8\_FLI  
US-10-425-114-23644

Query Match 0.9%; Score 20; DB 17; Length 1282;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 GCGCAACAGCAGCAGCAGC 59  
Db 562 GCGCAACAGCAGCAGCAGC 581

RESULT 72  
US-10-425-115-19062  
; Sequence 19062, Application US/10425115  
; Publication No. US20040214272A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53222)B  
; CURRENT APPLICATION NUMBER: US/10/425,115  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 369326  
; SEQ ID NO 19062

;  
; LENGTH: 1435  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: MRT4577\_117387C.1  
US-10-425-115-19062

Query Match 0.9%; Score 20; DB 18; Length 1435;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1908 CAACGAAGAAATGAAAA 1927  
Db 1218 CAACGAAGAAATGAAAA 1237

RESULT 73  
US-10-913-196-2/c  
; Sequence 2, Application US/10913196  
; Publication No. US20050053988A1  
; GENERAL INFORMATION:  
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS  
; APPLICANT: REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND  
; APPLICANT: HUMAN SERVICES  
; APPLICANT: Pastan, Ira H.  
; APPLICANT: Eglund, Kristi  
; APPLICANT: Lee, Byungkook  
; APPLICANT: Vincent, James  
; TITLE OF INVENTION: GENE EXPRESSED IN BREAST CANCER AND METHODS OF USE  
; FILE REFERENCE: 4239-66345-02  
; CURRENT APPLICATION NUMBER: US/10/913,196  
; CURRENT FILING DATE: 2004-08-05  
; PRIOR APPLICATION NUMBER: US 60/493,522  
; PRIOR FILING DATE: 2003-08-08  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2  
; LENGTH: 1439  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-913-196-2

Query Match 0.9%; Score 20; DB 19; Length 1439;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCAGCGGC 62  
Db 483 GCAACAGCAGCAGCAGCGGC 464

RESULT 74  
US-10-425-114-34740/c  
; Sequence 34740, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 34740  
; LENGTH: 1850  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:

```
; OTHER INFORMATION: Clone ID: UC-ZMFLMO17307H04_FLI
US-10-425-114-34740
Query Match      0.9%; Score 20; DB 17; Length 1850;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 47 CAGCAGCAGCAGCGGCTCCG 66
Db 485 CAGCAGCAGCAGCGGCTCCG 466
RESULT 75
US-10-310-154-280
; Sequence 280 Application US/10310154
; Publication No. US20030233670A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Chomet, Paul S.
; APPLICANT: Adams, Thomas H
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Agarwal, Ameeta K.
; APPLICANT: Ahrens, Jeffrey E.
; APPLICANT: Ball, James A.
; APPLICANT: Banu, G.
; APPLICANT: Bell, Erin
; APPLICANT: Boddupalli, Raghava
; APPLICANT: Deikman, Jill
; APPLICANT: Deng, Nolian
; APPLICANT: Dong, Jinzhuo
; APPLICANT: Duff, Stephen M.
; APPLICANT: Galligan, Meghan M.
; APPLICANT: Hinchey, Brenda S.
; APPLICANT: Huang, Shihshieh
; APPLICANT: Johnson, G. Richard
; APPLICANT: Jung, Vincent
; APPLICANT: Kretzmer, Keith A
; APPLICANT: Laccetti, Lucille B.
; APPLICANT: Lai, Chao-Qiang
; APPLICANT: Lee, Gary
; APPLICANT: Lin, Jier-Yi
; APPLICANT: Liu, Jingdong
; APPLICANT: Lu, Bin
; APPLICANT: Luethy, Michael M.
; APPLICANT: Lund, Adrian
; APPLICANT: Madson, Linda L.
; APPLICANT: Malloy, Kathleen A.
; APPLICANT: McKiel, Christine L.
; APPLICANT: Miller, Philip W.
; APPLICANT: Padmavathi, Manchikanti
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Start, William G.
; APPLICANT: Tennesen, Dan
; APPLICANT: Vidya, K.R.
; APPLICANT: Wang, Haiyun
; APPLICANT: Xin, Zhanqun
; APPLICANT: Xu, Nanfei
; APPLICANT: Yang, Chunzhi
; APPLICANT: Zeng, Xiaoping
; APPLICANT: Zhang, Qiang
; APPLICANT: Zhao, Yajuan
; APPLICANT: Zhou, Li
; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
; FILE REFERENCE: 38-15(52796)B
; CURRENT APPLICATION NUMBER: US/10/310,154
; CURRENT FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 60/337,358
; PRIOR FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 736
; SEQ ID NO 280
; LENGTH: 1891
; TYPE: DNA
; ORGANISM: Oryza sativa
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (193)...(1380)
; OTHER INFORMATION:
US-10-310-154-280
Query Match      0.9%; Score 20; DB 17; Length 1891;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 40 GCGGCAACACGACGACGACG 59
Db 1125 GCGGCAACACGACGACGACG 1144
Search completed: March 26, 2005, 01:14:24
Job time : 1240 secs
```

**This Page Blank (uspto)**

100

100

100

100

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2005, 17:12:50 ; Search time 408 Seconds  
(without alignments)  
9320.361 Million cell updates/sec

Title: US-09-436-184-3  
Perfect score: 2324  
Sequence: 1 cggaccgtgcaatggccag.....cttgggaactctggagaga 2324

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/prodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/prodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/prodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/prodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/prodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/prodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2324	100.0	2324	4	US-09-903-248-3
2	2324	100.0	2324	4	US-09-903-199-3
3	2324	100.0	2324	4	US-09-903-216-3
4	2324	100.0	2324	4	US-09-903-063-3
5	2324	100.0	2324	4	US-09-859-604-3
6	785	33.8	2442	3	US-09-040-485-1
7	640	27.5	2442	4	US-09-702-705-1791
8	640	27.5	2442	4	US-09-736-457-1791
9	640	27.5	2442	4	US-09-671-325-1791
10	640	27.5	2442	4	US-09-659-521-1
C 11	130	5.6	74730	4	US-09-948-016-15189
C 12	115	4.9	956	4	US-09-948-016-3447
C 13	88	3.8	5746	4	US-09-948-016-16453
C 14	77	3.3	601	4	US-09-948-016-122372
C 15	73	3.1	601	4	US-09-948-016-122387
C 16	59	2.5	2208	4	US-09-948-016-236
C 17	45	1.9	601	4	US-09-948-016-122345
C 18	22	0.9	9984	4	US-09-948-016-13283
C 19	21	0.9	1845	4	US-09-614-034-188
C 20	20	0.9	20	4	US-09-859-604-10
C 21	20	0.9	20	4	US-09-859-604-11
C 22	20	0.9	20	4	US-09-859-604-12
C 23	20	0.9	342	4	US-09-621-976-24
C 24	20	0.9	786	4	US-09-489-039A-1933
C 25	20	0.9	1071	4	US-09-948-016-136
C 26	20	0.9	1106	4	US-09-948-016-3254
C 27	20	0.9	1116	4	US-09-948-016-1812

28	20	0.9	8841	4	US-09-949-016-14996	Sequence 14996, A
29	20	0.9	11668	4	US-09-949-016-11878	Sequence 11878, A
30	20	0.9	11713	4	US-09-949-016-13554	Sequence 13554, A
31	20	0.9	47284	4	US-09-949-016-17029	Sequence 17029, A
C 32	20	0.9	117937	4	US-09-949-016-12762	Sequence 12762, A
C 33	20	0.9	117937	4	US-09-949-016-15775	Sequence 15775, A
C 34	19	0.8	148	4	US-09-513-999C-35281	Sequence 35281, A
C 35	19	0.8	309	4	US-09-248-796A-13462	Sequence 13462, A
C 36	19	0.8	601	4	US-09-949-016-61822	Sequence 61822, A
C 37	19	0.8	951	4	US-09-248-796A-4406	Sequence 4406, Ap
C 38	19	0.8	1599	4	US-09-949-016-1792	Sequence 1792, Ap
C 39	19	0.8	1623	3	US-08-339-214-25	Sequence 25, Appl
C 40	19	0.8	1751	3	US-09-149-476-110	Sequence 110, Appl
C 41	19	0.8	1821	4	US-09-949-016-5583	Sequence 5583, Ap
C 42	19	0.8	2286	4	US-09-620-312D-49	Sequence 49, Appl
C 43	19	0.8	2377	3	US-09-149-476-264	Sequence 264, Ap
C 44	19	0.8	3462	4	US-09-949-016-6642	Sequence 2642, Ap
C 45	19	0.8	3482	4	US-09-949-016-633	Sequence 633, Ap
C 46	19	0.8	6625	4	US-09-949-016-13534	Sequence 13534, A
C 47	19	0.8	7460	4	US-09-949-016-12375	Sequence 12375, A
C 48	19	0.8	7462	4	US-09-949-016-14384	Sequence 14384, A
C 49	19	0.8	7700	4	US-09-949-016-15619	Sequence 15619, A
C 50	19	0.8	28843	4	US-09-949-016-17325	Sequence 17325, A
C 51	19	0.8	300598	4	US-09-949-016-11868	Sequence 11868, A
C 52	19	0.8	302604	4	US-09-949-016-14588	Sequence 14588, A
C 53	19	0.8	302604	4	US-09-949-016-14589	Sequence 14589, A
C 54	19	0.8	308362	4	US-09-948-016-17119	Sequence 17119, A
C 55	19	0.8	828152	4	US-09-948-016-12777	Sequence 12777, A
C 56	18	0.8	22	4	US-09-859-604-13	Sequence 13, Appl
C 57	18	0.8	286	4	US-09-313-294A-3024	Sequence 3024, Ap
C 58	18	0.8	319	2	US-08-482-842B-2	Sequence 2, Appl
C 59	18	0.8	452	4	US-09-513-999C-10601	Sequence 10601, A
C 60	18	0.8	529	4	US-09-270-767-10408	Sequence 10408, A
C 61	18	0.8	557	4	US-09-918-686-4	Sequence 4, Appl
C 62	18	0.8	561	4	US-09-621-976-16896	Sequence 16896, A
C 63	18	0.8	601	4	US-09-949-016-36982	Sequence 36982, A
C 64	18	0.8	601	4	US-09-949-016-40425	Sequence 40425, A
C 65	18	0.8	601	4	US-09-949-016-42135	Sequence 42135, A
C 66	18	0.8	601	4	US-09-949-016-42136	Sequence 42136, A
C 67	18	0.8	601	4	US-09-949-016-42137	Sequence 42137, A
C 68	18	0.8	601	4	US-09-949-016-42138	Sequence 42138, A
C 69	18	0.8	601	4	US-09-948-016-42139	Sequence 42139, A
C 70	18	0.8	601	4	US-09-948-016-42140	Sequence 42140, A
C 71	18	0.8	601	4	US-09-949-016-53139	Sequence 53139, A
C 72	18	0.8	601	4	US-09-949-016-67303	Sequence 67303, A
C 73	18	0.8	601	4	US-09-949-016-69417	Sequence 69417, A
C 74	18	0.8	601	4	US-09-949-016-69418	Sequence 69418, A
C 75	18	0.8	601	4	US-09-949-016-69419	Sequence 69419, A
C 76	18	0.8	601	4	US-09-949-016-79910	Sequence 79910, A
C 77	18	0.8	601	4	US-09-949-016-79911	Sequence 79911, A
C 78	18	0.8	601	4	US-09-949-016-79912	Sequence 79912, A
C 79	18	0.8	601	4	US-09-949-016-80517	Sequence 80517, A
C 80	18	0.8	601	4	US-09-949-016-113358	Sequence 113358, A
C 81	18	0.8	601	4	US-09-949-016-113359	Sequence 113359, A
C 82	18	0.8	601	4	US-09-948-016-113360	Sequence 113360, A
C 83	18	0.8	601	4	US-09-948-016-153469	Sequence 153469, A
C 84	18	0.8	601	4	US-09-949-016-176839	Sequence 176839, A
C 85	18	0.8	601	4	US-09-949-016-177100	Sequence 177100, A
C 86	18	0.8	601	4	US-09-949-016-177101	Sequence 177101, A
C 87	18	0.8	601	4	US-09-949-016-177102	Sequence 177102, A
C 88	18	0.8	603	4	US-09-621-976-3826	Sequence 3826, Ap
C 89	18	0.8	615	4	US-09-513-999C-3903	Sequence 3903, Ap
C 90	18	0.8	625	4	US-09-270-767-13169	Sequence 13169, A
C 91	18	0.8	711	4	US-09-252-991A-2925	Sequence 2925, Ap
C 92	18	0.8	905	4	US-09-270-767-12936	Sequence 12936, A
C 93	18	0.8	934	3	US-09-174-65D-4	Sequence 4, Appl
C 94	18	0.8	934	3	US-09-595-564A-4	Sequence 4, Appl
C 95	18	0.8	958	3	US-09-457-046B-13	Sequence 13, Appl
C 96	18	0.8	958	3	US-09-866-570B-13	Sequence 13, Appl
C 97	18	0.8	988	4	US-09-949-016-2721	Sequence 2721, Ap
C 98	18	0.8	1014	2	US-09-066-074-1	Sequence 1, Appl
C 99	18	0.8	1014	2	US-08-555-912A-1	Sequence 1, Appl
C 100	18	0.8	1014	4	US-09-275-900-1	Sequence 1, Appl

## ALIGNMENTS

```
RESULT 1
US-09-903-248-3
; Sequence 3, Application US/09903248
; Patent No. 6783758
; GENERAL INFORMATION:
; APPLICANT: Wands, Jack R.
; APPLICANT: de la Monte, Suzanne M.
; APPLICANT: Ince, Nedim
; APPLICANT: Carlson, Rolf I.
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS
; FILE REFERENCE: 21486-032 DIVS
; CURRENT APPLICATION NUMBER: US/09/903,248
; CURRENT FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 09/436,184
; PRIOR FILING DATE: 1999-11-08
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2324
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-903-248-3

Query Match      100.0%; Score 2324; DB 4; Length 2324;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCGTGAATGGCCAGCGTAAGAAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60
Db 1 CGGACCGTGAATGGCCAGCGTAAGAAATGCCAAGAGCAGCGGCAACAGCAGCAGCAGCG 60

Qy 61 GCTCCGCGAGCGGTAGCAGCGAGTCCGGCGAGCAGCAGCCCGGGCCCGGAGAGACAA 120
Db 61 GCTCCGCGAGCGGTAGCAGCGAGTCCGGCGAGCAGCAGCCCGGGCCCGGAGAGACAA 120

Qy 121 AGCATGAGGACACAAGAAATGGGAGGAAAGCGGACTCTCGGGAACCTTCACTTCCAGT 180
Db 121 AGCATGAGGACACAAGAAATGGGAGGAAAGCGGACTCTCGGGAACCTTCACTTCCAGT 180

Qy 181 GGTTTATGGTGATTCGATTCGCGGCGTCTGGACATCTGTAGCTGTGCTTTGGTTTCATC 240
Db 181 GGTTTATGGTGATTCGATTCGCGGCGTCTGGACATCTGTAGCTGTGCTTTGGTTTCATC 240

Qy 241 TTGTTGACTATGAGGAAGTTCTAGGAAATCTAGGAATCTATGATGCTGATGGATGGAG 300
Db 241 TTGTTGACTATGAGGAAGTTCTAGGAAATCTAGGAATCTATGATGCTGATGGATGGAG 300

Qy 301 ATTTTGTATGGATGATGCCAAAGTTTATTTAGGACTTTAAGAGAGATCTACTTTCAGAGC 360
Db 301 ATTTTGTATGGATGATGCCAAAGTTTATTTAGGACTTTAAGAGAGATCTACTTTCAGAGC 360

Qy 361 CAGCAGTCCCGCCAGAGAGAGCTGAGCCACACACTGAGCCCGGAGGAGCAGTTTCTGTGG 420
Db 361 CAGCAGTCCCGCCAGAGAGAGCTGAGCCACACACTGAGCCCGGAGGAGCAGTTTCTGTGG 420

Qy 421 AGGCAGAACCCCGAATATCGAAGATGAGGAGAGAGACTTGCACAAAGAGATCGGCCACAG 480
Db 421 AGGCAGAACCCCGAATATCGAAGATGAGGAGAGAGACTTGCACAAAGAGATCGGCCACAG 480

Qy 481 AAATGGTACGCGAGACATGTTGAGGAGAGAGACTTGCACAAAGAGATCGGCCACAG 540
Db 481 AAATGGTACGCGAGACATGTTGAGGAGAGAGACTTGCACAAAGAGATCGGCCACAG 540

Qy 541 GAGAACCAACAAGAGGATGATGATTTCTTATGGGACTGTATGATGATGATGATGATG 600
Db 541 GAGAACCAACAAGAGGATGATGATTTCTTATGGGACTGTATGATGATGATGATGATG 600

Qy 601 AGACCCCTGGAAACCTGAAGTATCTCATGAGAAACCGGAGCATAGTTACCAGTGGAAAGAGA 660
```



1681 CATCTGTGCGAAGCTCCTACCTCAATGTAATGAGTGAAGACACAGCCTTGTGTGA 1740  
1741 CCCCAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGAGTTAA 1800  
1741 CCCCAAGAAACGGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGAGTTAA 1800  
1801 TCCGAGATGAAGGCTTGCAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1860  
1801 TCCGAGATGAAGGCTTGCAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1860  
1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTTCAAGCTGTGCGAGCAAGGAAGAA 1920  
1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTTCAAGCTGTGCGAGCAAGGAAGAA 1920  
1921 ATGAAATGCTGCAAGAGAGTCTTAAACCTGTACTAGAAAGTTCCCGGAGA 1980  
1921 ATGAAATGCTGCAAGAGAGTCTTAAACCTGTACTAGAAAGTTCCCGGAGA 1980  
1981 CAACAGATGCAGAGAGGACAGATCAATATTCATCATGCAACCCGGGACTCAGGTGT 2040  
1981 CAACAGATGCAGAGAGGACAGATCAATATTCATCATGCAACCCGGGACTCAGGTGT 2040  
2041 GCGCGCACAGAGGCCCCAACAATGCGAGCTCCGAATGCACTGGGCTTGGTGAATCCCA 2100  
2041 GCGCGCACAGAGGCCCCAACAATGCGAGCTCCGAATGCACTGGGCTTGGTGAATCCCA 2100  
2101 AGGAAGCTGCAAGATTCGATGTGCAACGAGACAGGACCTGGGAGGAGGCAAGGTGC 2160  
2101 AGGAAGCTGCAAGATTCGATGTGCAACGAGACAGGACCTGGGAGGAGGCAAGGTGC 2160  
2161 TCATCTTTGATGACTCTTTGAGCAGAGATGATGATGATGATGATGATGATGATGATGAT 2220  
2161 TCATCTTTGATGACTCTTTGAGCAGAGATGATGATGATGATGATGATGATGATGATGAT 2220  
2221 TATTTCATGCTGATGTGGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGG 2280  
2221 TATTTCATGCTGATGTGGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGG 2280  
2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2324  
2281 CAATTTAGCATGAATTCATGCAAGCTTGGGAACTCTGGAGAGA 2324

RESULT 2

US-09-903-199-3  
; Sequence 3, Application US/09903199  
; Patent No. 679696  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV4  
; CURRENT APPLICATION NUMBER: US/09/903,199  
; CURRENT FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-903-199-3

Query Match 100.0%; Score 2324; DB 4; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 CGGACCGTGCATGCGCCAGCGTGAAGATGCCAAGAGCAGCGGCAACGACGACGCG 60  
1 CGGACCGTGCATGCGCCAGCGTGAAGATGCCAAGAGCAGCGGCAACGACGACGCG 60

61 GCTCCGCGACGGGTAGCAGAGTGGCGGCGAGCAGCAGCCCGGGCCCGGAGAGACAA 120  
61 GCTCCGCGACGGGTAGCAGAGTGGCGGCGAGCAGCAGCCCGGGCCCGGAGAGACAA 120  
121 AGCATGAGAGACACAGAGATGGAGAAAGCGGAGCTCTCGGGAACCTTCACTTCCAGT 180  
121 AGCATGAGAGACACAGAGATGGAGAAAGCGGAGCTCTCGGGAACCTTCACTTCCAGT 180  
181 GGTATTATGGTGAATGTCATTGCTGGCGCTGCGACATCTGTAGCTGTCTGTTGGTTGATC 240  
181 GGTATTATGGTGAATGTCATTGCTGGCGCTGCGACATCTGTAGCTGTCTGTTGGTTGATC 240  
241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGTGATGGTGA 300  
241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGTGATGGTGA 300  
301 ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
301 ATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
361 CAGCAGTCCCGCGAGAGAGGCTGAGCCACACATGAGCCCGGAGGAGAGGTTCTTGTGG 420  
361 CAGCAGTCCCGCGAGAGAGGCTGAGCCACACATGAGCCCGGAGGAGAGGTTCTTGTGG 420  
421 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
421 AGGCAGAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
481 AAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540  
481 AAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540  
541 GAGAACCAACCAAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
541 GAGAACCAACCAAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 600  
601 AGACCTTGGAACTTCTCATGAGAAACCGAGCATAGTTACCACTGGAAGAGA 660  
601 AGACCTTGGAACTTCTCATGAGAAACCGAGCATAGTTACCACTGGAAGAGA 660  
661 CAGTTTCAAGAGCTGTAAATCAGGATATGGAAGAGATGATGCTGTGAGCAAGAAATCCAG 720  
661 CAGTTTCAAGAGCTGTAAATCAGGATATGGAAGAGATGATGCTGTGAGCAAGAAATCCAG 720  
721 ATTCAGTGAACCAAGTAGTAGAGATGAAAGATTCGACCATGATACAGATGATGATGATG 780  
721 ATTCAGTGAACCAAGTAGTAGAGATGAAAGATTCGACCATGATACAGATGATGATGATG 780  
781 ACCAAGTCTATGAGGAAACAGCAGATATGAACTCTAGAAATGAAAGGATGAAATCA 840  
781 ACCAAGTCTATGAGGAAACAGCAGATATGAACTCTAGAAATGAAAGGATGAAATCA 840  
841 CAGAAGTAACTGCTCCCTGAGGATAATCCTGTAGAGATTCACAGGTAAATTTGTAGAAG 900  
841 CAGAAGTAACTGCTCCCTGAGGATAATCCTGTAGAGATTCACAGGTAAATTTGTAGAAG 900  
901 AAGTAAGCATTTTCTGTGGAAGAACAGCAGAGAGTACCAACAGAAACAAATAGAAAAA 960  
901 AAGTAAGCATTTTCTGTGGAAGAACAGCAGAGAGTACCAACAGAAACAAATAGAAAAA 960  
961 CAGATGATCCAGAAACAAAGGTTAAGAAAGAGGCTTAACTTTTAAATTAAT 1020  
961 CAGATGATCCAGAAACAAAGGTTAAGAAAGAGGCTTAACTTTTAAATTAAT 1020  
1021 TTGATAAGACTATTAAAGCTGAATGATGATGATGATGATGATGATGATGATGATGATG 1080  
1021 TTGATAAGACTATTAAAGCTGAATGATGATGATGATGATGATGATGATGATGATGATG 1080  
1081 TTGAGGAAGCAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140  
1081 TTGAGGAAGCAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140

1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAAAGTAATGAGG 1200  
Db |||||  
1141 CAAGATATGGGAAGCGCAGTGTGAGGATGATTTGGCTGAGAGAGAGAAAGTAATGAGG 1200  
Qy |||||  
1201 TGCTACCTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCCTACCTGATGTCCTTCAG 1260  
Db |||||  
1201 TGCTACCTGGAGCCATCGAGACCTACCAAGAGGTGGCCAGCCTACCTGATGTCCTTCAG 1260  
Qy |||||  
1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAATTTCTAGGTCATATGA 1320  
Db |||||  
1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAATTTCTAGGTCATATGA 1320  
Qy |||||  
1321 GAGGTTCCCTGCTTACCCCTGCAGAGATTAGTTCAACTATTTCCTCAATGATACCTTCCTAA 1380  
Db |||||  
1321 GAGGTTCCCTGCTTACCCCTGCAGAGATTAGTTCAACTATTTCCTCAATGATACCTTCCTAA 1380  
Qy |||||  
1381 AAATGACCTTGGCGTGGGATACCTCTGTATAGAGATATGCAATGCAAGAAAGCTTT 1440  
Db |||||  
1381 AAATGACCTTGGCGTGGGATACCTCTGTATAGAGATATGCAATGCAAGAAAGCTTT 1440  
Qy |||||  
1441 ATGAAGAGGTGCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTATGGCTTCA 1500  
Db |||||  
1441 ATGAAGAGGTGCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATTATGGCTTCA 1500  
Qy |||||  
1501 TCCTGAAGGCACAGAAACAAATTCCTGAGAGCATCCCATATTAAAGAAAGGAATGAAT 1560  
Db |||||  
1501 TCCTGAAGGCACAGAAACAAATTCCTGAGAGCATCCCATATTAAAGAAAGGAATGAAT 1560  
Qy |||||  
1561 CCGGAGATCTGGCACTGATGATGGGAGATTTATTTCCACCTGGGGGATGCCATGCAGA 1620  
Db |||||  
1561 CCGGAGATCTGGCACTGATGATGGGAGATTTATTTCCACCTGGGGGATGCCATGCAGA 1620  
Qy |||||  
1621 GGGTTGGGAACAAAGAGGCATATAAGTGTATGAGCTTGGGCAACAGAGAGGACACTTTG 1680  
Db |||||  
1621 GGGTTGGGAACAAAGAGGCATATAAGTGTATGAGCTTGGGCAACAGAGAGGACACTTTG 1680  
Qy |||||  
1681 CATCTGCTGCAACAGCTCTACTTACAAATGTAATGAGTGAAGCAAGCAAGCTTTGGTGA 1740  
Db |||||  
1681 CATCTGCTGCAACAGCTCTACTTACAAATGTAATGAGTGAAGCAAGCAAGCTTTGGTGA 1740  
Qy |||||  
1741 CCCCAGAGAAACGGGCTACACAGATTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
Db |||||  
1741 CCCCAGAGAAACGGGCTACACAGATTAGTAAAGTCTTTAGAAAGAACTGGAAGTTAA 1800  
Qy |||||  
1801 TCCGAGATGAAGGCTTTGAGTGTATGATGAAGCAAGGCTCTCTCTGCTGAGGATG 1860  
Db |||||  
1801 TCCGAGATGAAGGCTTTGAGTGTATGATGAAGCAAGGCTCTCTCTGCTGAGGATG 1860  
Qy |||||  
1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTTCAAGCTGTGGCAGCAAGAAAGAA 1920  
Db |||||  
1861 AAAACCTGAGGAAAAAGGGGACTGGAGCCAGTTTCAAGCTGTGGCAGCAAGAAAGAA 1920  
Qy |||||  
1921 ATGAAGATGCTGCAAGAGGCTCTTAAACCTGTACTAGTAAAGTTTCCCGGAG 1980  
Db |||||  
1921 ATGAAGATGCTGCAAGAGGCTCTTAAACCTGTACTAGTAAAGTTTCCCGGAG 1980  
Qy |||||  
1981 CAACAGGATGAGAAAGGAGCAGATCAATATTCATCATGCAACCCCGGAGCTCACGCTG 2040  
Db |||||  
1981 CAACAGGATGAGAAAGGAGCAGATCAATATTCATCATGCAACCCCGGAGCTCACGCTG 2040  
Qy |||||  
2041 GGCGCGCACACAGGCGCCAAACTGCGAGGCTCCGAATGCACTGGGCTGTGATCCCA 2100  
Db |||||  
2041 GGCGCGCACACAGGCGCCAAACTGCGAGGCTCCGAATGCACTGGGCTGTGATCCCA 2100  
Qy |||||  
2101 AGGAAGCTGCAAGATTCGATGTCGCAACAGAGACCGAGGCTGGGAGGAGGCAAGGTGC 2160  
Db |||||  
2101 AGGAAGCTGCAAGATTCGATGTCGCAACAGAGACCGAGGCTGGGAGGAGGCAAGGTGC 2160  
Qy |||||  
2161 TCATCTTTGATGACTCTCTTTGAGCAGAGGATATGGCAGGATGCCTCATCTTTCCGGCTGA 2220  
Db |||||  
2161 TCATCTTTGATGACTCTCTTTGAGCAGAGGATATGGCAGGATGCCTCATCTTTCCGGCTGA 2220  
Qy |||||  
2221 TATTATGCTGGATGTGTGGCATCCGGAACTGACACACACAGCAGAGAGCGCCTTCCAG 2280

2221 TATTATGCTGGATGTGTGGCATCCGGAACCTGACACACAGCAGAGAGCGCTTCCAG 2280  
Qy |||||  
2281 CAATTTAGCATGAATTCATCAAGCTTGGAAAACTCTGGAGAGA 2324  
Db |||||  
2281 CAATTTAGCATGAATTCATCAAGCTTGGAAAACTCTGGAGAGA 2324  
Qy |||||  
RESULT 3  
US-09-903-216-3  
; Sequence 3, Application US/09903216  
; Patent No. 6812206  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M.  
; APPLICANT: Ince, Nedim  
; APPLICANT: Carlson, Rolf I.  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 DIV2  
; CURRENT APPLICATION NUMBER: US/09/903,216  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-903-216-3  
Query Match 100.0%; Score 2324; DB 4; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 CGGACCGTGCATATGGCCAGCGTTAAGATGCCAAGAGCAGCGGCCCAACAGCAGCAGCAGCG 60  
Db 1 CGGACCGTGCATATGGCCAGCGTTAAGATGCCAAGAGCAGCGGCCCAACAGCAGCAGCAGCG 60  
Qy 61 GCTCCGCGCAGCGGTAGCACAGTGGCGGAGCAGCAGCCCCGGGCCCCGGAGAGAGACAA 120  
Db 61 GCTCCGCGCAGCGGTAGCACAGTGGCGGAGCAGCAGCCCCGGGCCCCGGAGAGAGACAA 120  
Qy 121 AGCATGGAGAGACAAAGAAATGGGAGAAAGCGGAGCTCTCGGAACTTCATTTCTTACGT 180  
Db 121 AGCATGGAGAGACAAAGAAATGGGAGAAAGCGGAGCTCTCGGAACTTCATTTCTTACGT 180  
Qy 181 GGTATTGCTGATGCTGCTGGCGCTCTGGACATCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 240  
Db 181 GGTATTGCTGATGCTGCTGGCGCTCTGGACATCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 240  
Qy 241 TTGTTGACTATGAGGAAGTTCTAGGAAAACTAGGAATCTATGATGCTGATGCTGATGCTGATG 300  
Db 241 TTGTTGACTATGAGGAAGTTCTAGGAAAACTAGGAATCTATGATGCTGATGCTGATGCTGATG 300  
Qy 301 ATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
Db 301 ATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360  
Qy 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGCTTCTCTGG 420  
Db 361 CAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCCGAGGAGCAGCTTCTCTGG 420  
Qy 421 AGGCAGAAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
Db 421 AGGCAGAAACCCCAAGATATCGAAGATGAAGCAAAAGAAACAAATTCAGTCCCTTCTCCATG 480  
Qy 481 AAATGGTACACGAGAACATGTTGAGGAGAGAGACTTGGCAACAGAGAGATGGACCCACAG 540  
Db 481 AAATGGTACACGAGAACATGTTGAGGAGAGAGACTTGGCAACAGAGAGATGGACCCACAG 540  
Qy 541 GAGAACCAACAGAGAGGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600

Db 541 GAGAACCAACAGAGGATGATGAGTTTCTTATGGCAGCTGATGTAGATGATGATTG 600  
 Qy 601 AGACCTGGAACCTGAAGTATCTCATGAAGAACCGAGATAGTTTACCATGTGAGAGAGA 660  
 Db 601 AGACCTGGAACCTGAAGTATCTCATGAAGAACCGAGATAGTTTACCATGTGAGAGAGA 660  
 Qy 661 CAGTTTCAAGAGCTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720  
 Db 661 CAGTTTCAAGAGCTGTAATCAGGATATGGAAGAGATGATGCTGAGCAGGAAATCCAG 720  
 Qy 721 ATTCCAGTGAACCAAGTATGAGATGAAAGATTGACCATGATACAGATGATGTAACAT 780  
 Db 721 ATTCCAGTGAACCAAGTATGAGATGAAAGATTGACCATGATACAGATGATGTAACAT 780  
 Qy 781 ACCAAGTCTATGAGGAACAAGCAGTATATGAACTCTAGAAATGAAAGGATAGAAATCA 840  
 Db 781 ACCAAGTCTATGAGGAACAAGCAGTATATGAACTCTAGAAATGAAAGGATAGAAATCA 840  
 Qy 841 CAGAAGTAACTGCTCCCTGAGGATATCTCTGTAGAGATATTCACAGTAAATGTAAG 900  
 Db 841 CAGAAGTAACTGCTCCCTGAGGATATCTCTGTAGAGATATTCACAGTAAATGTAAG 900  
 Qy 901 AAGTAAGCATTTTTCTGTGGAAGAACAGCAGGAAAGTACCACAGAAACAAATAGAAAA 960  
 Db 901 AAGTAAGCATTTTTCTGTGGAAGAACAGCAGGAAAGTACCACAGAAACAAATAGAAAA 960  
 Qy 961 CAGATGATCCAGAAACAAAAGCAAAAGTTAAGAAAAGAACCTAAATCTTTAAATAAAT 1020  
 Db 961 CAGATGATCCAGAAACAAAAGCAAAAGTTAAGAAAAGAACCTAAATCTTTAAATAAAT 1020  
 Qy 1021 TTGATAAGACTATTAAGCTGAACCTGATGCTGCAGAAAACCTCGTAAAGGGGAAAA 1080  
 Db 1021 TTGATAAGACTATTAAGCTGAACCTGATGCTGCAGAAAACCTCGTAAAGGGGAAAA 1080  
 Qy 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140  
 Db 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTACGCAAAATACCTCAGAGTCCACGAG 1140  
 Qy 1141 CAAAGATATGGGAAGCCGAGTGTGAGGATGATTGGCTGAGAGAGAGAAAGTAATGAGG 1200  
 Db 1141 CAAAGATATGGGAAGCCGAGTGTGAGGATGATTGGCTGAGAGAGAGAAAGTAATGAGG 1200  
 Qy 1201 TGCTAGTGGAGCCATCGAGACCTACCAAGAGGTGGCAGCCTACCTGATGCTCCGCGAG 1260  
 Db 1201 TGCTAGTGGAGCCATCGAGACCTACCAAGAGGTGGCAGCCTACCTGATGCTCCGCGAG 1260  
 Qy 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTCTCAGACAGGCAACAAATTTCTAGGTCAATGA 1320  
 Db 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTCTCAGACAGGCAACAAATTTCTAGGTCAATGA 1320  
 Qy 1321 GAGGTTCCTCTTACCCTGAGAGATTAGTTCAACTATTTCCCAATGATACTTCTCTAA 1380  
 Db 1321 GAGGTTCCTCTTACCCTGAGAGATTAGTTCAACTATTTCCCAATGATACTTCTCTAA 1380  
 Qy 1381 AAAATGACCTTGGCGTGGGATACCTCTGATAGGAGATAATGACAATGCAAGAAAGTTT 1440  
 Db 1381 AAAATGACCTTGGCGTGGGATACCTCTGATAGGAGATAATGACAATGCAAGAAAGTTT 1440  
 Qy 1441 ATGAAGAGTCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATATGCGTTCA 1500  
 Db 1441 ATGAAGAGTCTGAGTGTGACACCTAATGATGGCTTTGCTAAAGTCCATATGCGTTCA 1500  
 Qy 1501 TCCTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCCATATTTAAAGGAAAGGAATAGAAT 1560  
 Db 1501 TCCTGAAGGCACAGAAACAAATTTGCTGAGAGCATCCCATATTTAAAGGAAAGGAATAGAAT 1560  
 Qy 1561 CCGGAGATCTTGGCACTGATGATGGAGATTTTATTTCCACTCGGGGATGCCATCGACA 1620  
 Db 1561 CCGGAGATCTTGGCACTGATGATGGAGATTTTATTTCCACTCGGGGATGCCATCGACA 1620  
 Qy 1621 GGGTTGGGAAACAAAGAGGCATATAAGTGTATGAGCTTGGGCAACAGAGGACACTTTG 1680  
 Db 1621 GGGTTGGGAAACAAAGAGGCATATAAGTGTATGAGCTTGGGCAACAGAGGACACTTTG 1680

Qy 1681 CATCTGTCTGGCAACGCTCACTCTACATGTGAATGAGCTGAAGCAGACGCTTGGTGA 1740  
 Db 1681 CATCTGTCTGGCAACGCTCACTCTACATGTGAATGAGCTGAAGCAGACGCTTGGTGA 1740  
 Qy 1741 CCCCCAAAAGAACCGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAAGTTAA 1800  
 Db 1741 CCCCCAAAAGAACCGGCTACACAGAGTTAGTAAAGTCTTTAGAAAGAACTGGAAAGTTAA 1800  
 Qy 1801 TCCGAGATGAAGGCTTTGCGAGTGTATGATAAAGCCAAAGGCTCTTCTCTGCTCAGGATG 1860  
 Db 1801 TCCGAGATGAAGGCTTTGCGAGTGTATGATAAAGCCAAAGGCTCTTCTCTGCTCAGGATG 1860  
 Qy 1861 AAAACCTCAGGGAAGAAAGGGGACTGGAGCCAGTTCACGCTGGCAGCAGGAGAGAGAA 1920  
 Db 1861 AAAACCTCAGGGAAGAAAGGGGACTGGAGCCAGTTCACGCTGGCAGCAGGAGAGAGAA 1920  
 Qy 1921 ATGAAAATGCTCTCAAAAGGAGCTCTTAAACCTGTACTTACTAGAAAAGTTCCCGAGA 1980  
 Db 1921 ATGAAAATGCTCTCAAAAGGAGCTCTTAAACCTGTACTTACTAGAAAAGTTCCCGAGA 1980  
 Qy 1981 CAAAGAGTGCAGAGAGAGAGAGATCAAAATATTTCCATCATGCAACCCCGGAGCTCACGTGT 2040  
 Db 1981 CAAAGAGTGCAGAGAGAGAGAGATCAAAATATTTCCATCATGCAACCCCGGAGCTCACGTGT 2040  
 Qy 2041 GSCCCACACAGAGGCCACAAACTGCAGGCTCCGAATGCACCTGGGCTTGGTGAATCCCA 2100  
 Db 2041 GSCCCACACAGAGGCCACAAACTGCAGGCTCCGAATGCACCTGGGCTTGGTGAATCCCA 2100  
 Qy 2101 AGGAAGGCTCAAGATTCGATGTGCCAACAGAGACAGAGACCTGGGAGAGAGGCAAGTGC 2160  
 Db 2101 AGGAAGGCTCAAGATTCGATGTGCCAACAGAGACAGAGACCTGGGAGAGAGGCAAGTGC 2160  
 Qy 2161 TCATCTTTGTGACTCTTTTGAGCAGAGGATGAGCAGATGCCTCATCTTTCCGGCTGA 2220  
 Db 2161 TCATCTTTGTGACTCTTTTGAGCAGAGGATGAGCAGATGCCTCATCTTTCCGGCTGA 2220  
 Qy 2221 TATTCATCGTGGATGTGGCTCCGAACTCGAACACAGAGAGAGAGAGAGAGAGAGAG 2280  
 Db 2221 TATTCATCGTGGATGTGGCTCCGAACTCGAACACAGAGAGAGAGAGAGAGAGAGAG 2280  
 Qy 2281 CAATTTAGCATGAATTCATCGAGCTTGGGAACTCTGGAGAGA 2324  
 Db 2281 CAATTTAGCATGAATTCATCGAGCTTGGGAACTCTGGAGAGA 2324

RESULT 4

US-09-903-063-3  
 ; Sequence 3, Application US/09903063  
 ; Patent No. 6815415  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wands, Jack R.  
 ; APPLICANT: de la Monte, Suzanne M.  
 ; APPLICANT: Ince, Nedim  
 ; APPLICANT: Carlson, Rolf I.  
 ; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
 ; FILE REFERENCE: 21486-032 DIV3  
 ; CURRENT APPLICATION NUMBER: US/09/903,063  
 ; CURRENT FILING DATE: 2001-10-11  
 ; PRIOR APPLICATION NUMBER: 09/436,184  
 ; NUMBER OF SEQ ID NOS: 9  
 ; SOFTWARE: Patent In Ver. 2.1  
 ; SEQ ID NO 3  
 ; LENGTH: 2324  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-903-063-3

Query Match 100.0%; Score 2324; DB 4; Length 2324;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Db 2161 TCATCTTTGATGACTCCTTTGAGCAGGAGGTATGCGAGGATGCGCTCATCTTTCCGGCTGA 2220  
Qy 2221 TATTCATCGTGGATGTGTGGCATCCGGAACCTGACACACACAGCAGACGACGCTTCCAG 2280  
Db 2221 TATTCATCGTGGATGTGTGGCATCCGGAACCTGACACACACAGCAGACGACGCTTCCAG 2280  
Qy 2281 CAATTTAGCATGAATTCATGCAAGCTTTGGAAACTCTGGAGAGA 2324  
Db 2281 CAATTTAGCATGAATTCATGCAAGCTTTGGAAACTCTGGAGAGA 2324

## RESULT 5

US-09-859-604-3  
; Sequence 3, Application US/09859604  
; Patent No. 6835370  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M  
; APPLICANT: Deutch, Alan H  
; APPLICANT: Ghanbari, Hossein A  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 CIP  
; CURRENT APPLICATION NUMBER: US/09/859,604  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 2324  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-859-604-3

Query Match 100.0%; Score 2324; DB 4; Length 2324;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGACCGTCAATGGCCCGCAGTGAAGATGCCAAGACAGCGCGCAACAGCAGCAGCAGCG 60  
Db 1 CGGACCGTCAATGGCCCGCAGTGAAGATGCCAAGACAGCGCGCGCAACAGCAGCAGCAGCG 60  
Qy 61 GCTCCGCGAGCGGTAGCAGCAGGTGCGGGCAGCAGCAGCGCCCGGGCCCGGAGAGACAA 120  
Db 61 GCTCCGCGAGCGGTAGCAGCAGGTGCGGGCAGCAGCAGCGCCCGGGCCCGGAGAGACAA 120  
Qy 121 AGCATGGAGGACACAGAATGGGAGGAAAGCGGAGCTCTCGGGAACTTCAATTTCCACGT 180  
Db 121 AGCATGGAGGACACAGAATGGGAGGAAAGCGGAGCTCTCGGGAACTTCAATTTCCACGT 180  
Qy 181 GGTATTATGGTATGATGCTGGGGCTCTGGACATCTGTAGCTGTGCTTTGGTTTGTATC 240  
Db 181 GGTATTATGGTATGATGCTGGGGCTCTGGACATCTGTAGCTGTGCTTTGGTTTGTATC 240  
Qy 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGTGATGGATGGAG 300  
Db 241 TTGTTGACTATGAGGAAGTTCTAGGAAACTAGGAATCTATGATGCTGTGATGGATGGAG 300  
Qy 301 ATTTGATGTGGATGATGCGAAATTTTATAGGCTTAAAGAGAGATCTACTTCAGAGC 360  
Db 301 ATTTGATGTGGATGATGCGAAATTTTATAGGCTTAAAGAGAGATCTACTTCAGAGC 360  
Qy 361 CAGCAGTCCCGCCAGAGAGCGCTGAGCCACACACTGAGCCCGGAGGAGCGTTCTGTGG 420  
Db 361 CAGCAGTCCCGCCAGAGAGCGCTGAGCCACACACTGAGCCCGGAGGAGCGTTCTGTGG 420  
Qy 421 AGGCGAGACCCCAAGTATCGAAGATGGAAGAAAGACAAATTCAGTCCCTTCTCCATG 480  
Db 421 AGGCGAGACCCCAAGTATCGAAGATGGAAGAAAGACAAATTCAGTCCCTTCTCCATG 480  
Qy 481 AAATGGTACACGCGAGAACATGTTGAGGGAGAGAGCTTGCACCAAGAGAGTGGACCCACAG 540  
Db 481 AAATGGTACACGCGAGAACATGTTGAGGGAGAGAGCTTGCACCAAGAGAGTGGACCCACAG 540

Qy 541 GAGAACCAACCAAGAGGATGATGATTTCTTTATGCGGACGTGATGTAGATGATGATTTG 600  
Db 541 GAGAACCAACCAAGAGGATGATGATTTCTTTATGCGGACGTGATGTAGATGATGATTTG 600  
Qy 601 AGACCTCGGAACCTGGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGCTGGAAGAGA 660  
Db 601 AGACCTCGGAACCTGGAAGTATCTCATGAAGAAACCGAGCATAGTTACACGCTGGAAGAGA 660  
Qy 661 CAGTTTCACAAGACTGTGTAATCAGGATATGGAAGAGATGATGCTGTGACGAGAAATCCAG 720  
Db 661 CAGTTTCACAAGACTGTGTAATCAGGATATGGAAGAGATGATGCTGTGACGAGAAATCCAG 720  
Qy 721 ATTCAGTGAACCAAGTAGTAGAGATGGAAGATTTGCCACATGATACACAGATGATGTACAT 780  
Db 721 ATTCAGTGAACCAAGTAGTAGAGATGGAAGATTTGCCACATGATACACAGATGATGTACAT 780  
Qy 781 ACCAAGTCTATGAGGAAACCAAGCATATATGAACCTCTAGAAAATGAAGGATGAGAAATCA 840  
Db 781 ACCAAGTCTATGAGGAAACCAAGCATATATGAACCTCTAGAAAATGAAGGATGAGAAATCA 840  
Qy 841 CAGAAGTAACCTGCTCCCTCCCTGAGGATTAATCTCTGTAGAAGATTCACAGGTAAATTTGAGAAG 900  
Db 841 CAGAAGTAACCTGCTCCCTCCCTGAGGATTAATCTCTGTAGAAGATTCACAGGTAAATTTGAGAAG 900  
Qy 901 AAGTAAGCATTTTCTCTGTGGAAGAACAGCAGGAGTACCAACAGGAAATTAAGTAAAT 1020  
Db 901 AAGTAAGCATTTTCTCTGTGGAAGAACAGCAGGAGTACCAACAGGAAATTAAGTAAAT 1020  
Qy 1021 TTGTAAGACTATTAAAGCTGAACTTGAATGCTGCGAGAAAAAATCCGTAAGGGGAAAAA 1080  
Db 1021 TTGTAAGACTATTAAAGCTGAACTTGAATGCTGCGAGAAAAAATCCGTAAGGGGAAAAA 1080  
Qy 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGCAGCAAAATACCTTCAGAGTCCACGAG 1140  
Db 1081 TTGAGGAAGCAGTGAATGCAATTTAAAGAACTAGTAGTAGCAGCAAAATACCTTCAGAGTCCACGAG 1140  
Qy 1141 CAGATATGGAAGGGCGAGTGTGAGATGATTTGGCTGCAAGAGGAGAGTAATGAGG 1200  
Db 1141 CAGATATGGAAGGGCGAGTGTGAGATGATTTGGCTGCAAGAGGAGAGTAATGAGG 1200  
Qy 1201 TGCTACGTGAGGCGCATCGAGACCTACCAAGAGGTGGCCAGCTACCTGATGTCCTGCGAG 1260  
Db 1201 TGCTACGTGAGGCGCATCGAGACCTACCAAGAGGTGGCCAGCTACCTGATGTCCTGCGAG 1260  
Qy 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAAATTTCTAGGTCAATGA 1320  
Db 1261 ACCTGCTGAAGCTGAGTTTGAAGCGTGCCTCAGACAGGCAACAAATTTCTAGGTCAATGA 1320  
Qy 1321 GAGGTTCCTGCTTACCTCGAGAGATTAGTTCAACTATTTCCCAATGATCTTCCCTTAA 1380  
Db 1321 GAGGTTCCTGCTTACCTCGAGAGATTAGTTCAACTATTTCCCAATGATCTTCCCTTAA 1380  
Qy 1381 AAAATGACCTTGGCGTGGGATACCTCTGATAGGAGTAATGCAATGCAAGAAAGATTT 1440  
Db 1381 AAAATGACCTTGGCGTGGGATACCTCTGATAGGAGTAATGCAATGCAAGAAAGATTT 1440  
Qy 1441 ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGGCTTTGCTAAAGTCCATTTATGGCTTCA 1500  
Db 1441 ATGAAGAGGTGCTGAGTGTGACACCTTAATGATGGCTTTGCTAAAGTCCATTTATGGCTTCA 1500  
Qy 1501 TCCTGAAGGACACAGAAACAAATTTGCTGAGAGCATCCCATATTTAAAGAGGAAATAGAAT 1560  
Db 1501 TCCTGAAGGACACAGAAACAAATTTGCTGAGAGCATCCCATATTTAAAGAGGAAATAGAAT 1560  
Qy 1561 CGGAGATCTCGGACATGATGGGAGATTTTATTTCCACCTGGGGATGCCATGCGAGA 1620  
Db 1561 CGGAGATCTCGGACATGATGGGAGATTTTATTTTCACTGGGGATGCCATGCGAGA 1620









Db 406 CAAGAAGATGACCCACAGGAGAACCAACAAGAGGATGATGAGTTCTTTATGGCGACT 465  
QY 582 GATGTAGATGATAGATTGAGACCCCTGGAACCTGAAGTATCTCATGAGGAACCGAGCAT 641  
Db 466 GATGTAGATGATAGATTGAGACCCCTGGAACCTGGAAGTATCTCATGAGGAACCGAGCAT 525  
QY 642 AGTTACCAAGTGGAGAGACAGTTTCAACAAGTCTGTAATCAGGATATGGAAGAGATGATG 701  
Db 526 AGTTACCAAGTGGAGAGACAGTTTCAACAAGTCTGTAATCAGGATATGGAAGAGATGATG 585  
QY 702 TCTGAGCAGGAAAATCCAGATTCAGTGAACCAAGTATGAGAGATGAAAGATTGCAACCAT 761  
Db 586 TCTGAGCAGGAAAATCCAGATTCAGTGAACCAAGTATGAGAGATGAAAGATTGCAACCAT 645  
QY 762 GATACAGATGATGTAACATACCAGTCTATGAGGAACCAAGTATGAGAGATGAAAGATTGCAACCAT 821  
Db 646 GATACAGATGATGTAACATACCAGTCTATGAGGAACCAAGTATGAGAGATGAAAGATTGCAACCAT 705  
QY 822 AATGAAGGGATAGAAAATCACAGAAGTAACTG 852  
Db 706 AATGAAGGGATAGAAAATCACAGAAGTAACTG 736

## RESULT 9

US-09-671-325-1791

; Sequence 1791, Application US/09671325

; Patent No. 6667154

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Lodes, Michael A.

; APPLICANT: Fanger, Gary

; APPLICANT: Vedwick, Tom

; APPLICANT: Carter, Darrick

; APPLICANT: Retter, Marc

; APPLICANT: Mannion, Jane

; APPLICANT: Fan, Liqun

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER

; CURRENT APPLICATION NUMBER: US/09/671,325

; CURRENT FILING DATE: 2000-09-26

; NUMBER OF SEQ ID NOS: 1825

; SOFTWARE: Fast-Seq for Windows Version 3.0

; SEQ ID NO 1791

; LENGTH: 2442

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-671-325-1791

Query Match 27.5%; Score 640; DB 4; Length 2442;  
Best Local Similarity 99.9%; Pred. No. 3e-297;  
Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 162 GGAACCTTCATCTTCCAGCTGGTTTATGGTGTATGCTGCTGGCGCTGCGACATCTGTA 221  
Db 46 GGAACCTTCATCTTCCAGCTGGTTTATGGTGTATGCTGCTGGCGCTGCGACATCTGTA 105  
QY 222 GCTGCTGGTTTGGTTTGTATGCTGATGAGGAGTTCTAGGAAACTAGGAATCTAT 281  
Db 106 GCTGCTGGTTTGGTTTGTATGCTGATGAGGAGTTCTAGGAAACTAGGAATCTAT 165  
QY 282 GATGCTGATGTTGATGAGGAGTTTTCATGCTGATGATGCAAGTTTATAGGACTTAAA 341  
Db 166 GATGCTGATGTTGATGAGGAGTTTTCATGCTGATGATGCAAGTTTATAGGACTTAAA 225  
QY 342 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCC 401  
Db 226 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCC 285  
QY 402 GAGAGCAGGTTCTTGTGGAGGCAACCCCAAGATATCGAAGATGAAGCAAGAACAA 461

Db 286 GAGGAGCAGGTTCTGTGGAGCGAAGACCCAGAAATATCGAAGATGAAGCAAAAGAACAA 345  
QY 462 ATTCAAGTCCCTTCTCCATGAAATGCTACACGACGAAACATGTTGAGGGAGAGACTTGCAG 521  
Db 346 ATTCAAGTCCCTTCTCCATGAAATGCTACACGACGAAACATGTTGAGGGAGAGACTTGCAG 405  
QY 522 CAAGAAGATGGAACCCACAGGAGAACCAACAAGAGGATGATGAGTTTCTTTATGGCGACT 581  
Db 406 CAAGAAGATGGAACCCACAGGAGAACCAACAAGAGGATGATGAGTTTCTTTATGGCGACT 465  
QY 582 GATGTAGATGATAGATTGAGACCCCTGGAACCTGAACTATCTCATGAGGAACCGAGCAT 641  
Db 466 GATGTAGATGATAGATTGAGACCCCTGGAACCTGAACTATCTCATGAGGAACCGAGCAT 525  
QY 642 AGTTACCAAGTGGAGAGACAGTTTCAACAAGTCTGTAATCAGGATATGGAAGAGATGATG 701  
Db 526 AGTTACCAAGTGGAGAGACAGTTTCAACAAGTCTGTAATCAGGATATGGAAGAGATGATG 585  
QY 702 TCTGAGCAGGAAAATCCAGATTCAGTGAACCAAGTATGAGAGATGAAAGATTGCAACCAT 761  
Db 586 TCTGAGCAGGAAAATCCAGATTCAGTGAACCAAGTATGAGAGATGAAAGATTGCAACCAT 645  
QY 762 GATACAGATGATGTAACATACCAGTCTATGAGGAACCAAGTATGAGAGATGAAAGATTGCAACCAT 821  
Db 646 GATACAGATGATGTAACATACCAGTCTATGAGGAACCAAGTATGAGAGATGAAAGATTGCAACCAT 705  
QY 822 AATGAAGGGATAGAAAATCACAGAAGTAACTG 852  
Db 706 AATGAAGGGATAGAAAATCACAGAAGTAACTG 736

## RESULT 10

US-09-659-521-1

; Sequence 1, Application US/09659521

; Patent No. 6727080

; GENERAL INFORMATION:

; APPLICANT: RADOSEVICH, JAMES A.

; TITLE OF INVENTION: GENE ENCODING LABYRINTHIN, A MARKER FOR CANCER

; FILE REFERENCE: 21511/91057

; CURRENT APPLICATION NUMBER: US/09/659,521

; CURRENT FILING DATE: 2000-09-12

; PRIOR APPLICATION NUMBER: PCT/US99/05365

; PRIOR FILING DATE: 1999-03-11

; PRIOR APPLICATION NUMBER: 09/040,485

; PRIOR FILING DATE: 1998-03-17

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 2442

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (70)..(834)

US-09-659-521-1

Query Match 27.5%; Score 640; DB 4; Length 2442;  
Best Local Similarity 99.9%; Pred. No. 3e-297;  
Matches 690; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 162 GGAACCTTCATCTTCCAGCTGGTTTATGGTGTATGCTGCTGGCGCTGCGACATCTGTA 221  
Db 46 GGAACCTTCATCTTCCAGCTGGTTTATGGTGTATGCTGCTGGCGCTGCGACATCTGTA 105  
QY 222 GCTGCTGGTTTGGTTTGTATGCTGATGAGGAGTTCTAGGAAACTAGGAATCTAT 281  
Db 106 GCTGCTGGTTTGGTTTGTATGCTGATGAGGAGTTCTAGGAAACTAGGAATCTAT 165  
QY 282 GATGCTGATGTTGATGAGGAGTTTTCATGCTGATGATGCAAGTTTATAGGACTTAAA 341  
Db 166 GATGCTGATGTTGATGAGGAGTTTTCATGCTGATGATGCAAGTTTATAGGACTTAAA 225  
QY 342 GAGAGATCTACTTCAGAGCCAGCAGTCCCGCAGAGAGGCTGAGCCACACACTGAGCCC 401

Db	226	GAGAGATCTATTTCAGAGCCAGCAGTCCCGCCGAGGAAGGGCTGAGCCACACTGAGCCC	285
Qy	402	GAGGAGCAGGTTCTCTGTGGAGGAGAAACCCACAGATATTCGAAGATGAAGCAAGAACA	461
Db	286	GAGGAGCAGGTTCTCTGTGGAGGAGAAACCCACAGATATTCGAAGATGAAGCAAGAACA	345
Qy	462	ATTGAGTCCCTTCTCCATGAAATGGTACGCGAGAACATGTTGAGGAGGAAGACTTTGCCAA	521
Db	346	ATTGAGTCCCTTCTCCATGAAATGGTACGCGAGAACATGTTGAGGAGGAAGACTTTGCCAA	405
Qy	522	CAAGAAGATGACCCACAGGAGAAACCAACAAGAGAGGATGATGAGTTTCTTATGGCGACT	581
Db	406	CAAGAAGATGACCCACAGGAGAAACCAACAAGAGGATGATGAGTTTCTTATGGCGACT	465
Qy	582	GATGTAGATGATAGATTGTGAGCCCTGGAAACCTGGAAGTATCTCATGAAGAAACCGAGCAT	641
Db	466	GATGTAGATGATAGATTGTGAGCCCTGGAACTTGAAGTATCTCATGAAGAAACCGAGCAT	525
Qy	642	AGTTACACAGTGGAGAGACAGTTTCACAGAGACTCTAATCAGGATATCGAAGAGATGATG	701
Db	526	AGTTACACAGTGGAGAGACAGTTTCACAGAGACTCTAATCAGGATATCGAAGAGATGATG	585
Qy	702	TCTGAGCAGGAAAATCAGATTCCAGTGAAACAGTGTAGTGAAGATGAAGATTGCAACAT	761
Db	586	TCTGAGCAGGAAAATCAGATTCCAGTGAAACAGTGTAGTGAAGATGAAGATTGCAACAT	645
Qy	762	GATCAGATGATGTAACATACCAAGTCTTATGAGGAAACAAGCAGTATATGAACCTCTAGAA	821
Db	646	GATCAGATGATGTAACATACCAAGTCTTATGAGGAAACAAGCAGTATATGAACCTCTAGAA	705
Qy	822	AATGAAGGATAGAAATCACAAGTAACTG	852
Db	706	AATGAAGGATAGAAATCACAAGTAACTG	736

```

RESULT 11
US-09-949-016-15189/c
; Sequence 15189, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15189
; LENGTH: 74730
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(74730)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15189

```

Qy	561	GATGAGTTCTTTATGGCGACTGATGTAGATGAGTTTGTAGACCCCTGGAACTGGAAGTA	620
Db	29846	GATGAGTTCTTTATGGCGACTGATGTAGATGAGTTTGTAGACCCCTGGAACTGGAAGTA	29787
Qy	621	TCTCATGAAG	630
Db	29786	TCTCATGAAG	29777

RESULT 12  
 US-09-949-016-3447/c  
 ; Sequence 3447, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; TITLE OF INVENTION: WITH A HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 3447  
 ; LENGTH: 956  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-09-949-016-3447

Query Match	4.9%	Score 115;	DB 4;	Length 956;
Best Local Similarity	100.0%;	Pred. No. 6.4e-45;		
Matches 115;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

Qy	516	TTGCAACAAGAAGATGGACCCACAGGAGAACCAACAAGAGATGATGAGTTCTTTATG	575
Db	462	TTGCAACAAGAAGATGGACCCACAGGAGAACCAACAAGAGATGATGAGTTCTTTATG	403
Qy	576	CGCGACTGATGTAGATGAGTTTGTAGACCCCTGGAACTGGAAGTATCTCATGAAG	630
Db	402	CGCGACTGATGTAGATGAGTTTGTAGACCCCTGGAACTGGAAGTATCTCATGAAG	348

```

RESULT 13
US-09-949-016-16453/c
? Sequence 16453, Application US/09949016
? Patent No. 6812339
? GENERAL INFORMATION:
? APPLICANT: VENTER, J. Craig et al.
? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: CU001307
? CURRENT APPLICATION NUMBER: US/09/949,016
? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO:16453
? LENGTH: 5746
? TYPE: DNA
? ORGANISM: Human
US-09-949-016-16453
Query Match 3.8%; Score 88; DB 4; Length 5746;

```

Best Local Similarity 100.0%; Pred. No. 6e-32;  
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1073 GGGAAAATTCAGGAGCAGTGATGCAATTTAAAGAACTAGTAGCAGAAATACCCCTCAGAG 1132  
Db 696 GGGAAAATTCAGGAGCAGTGATGCAATTTAAAGAACTAGTAGCAGAAATACCCCTCAGAG 637  
Qy 1133 TCCACGAGCAAGATATGGGAGGCGCAG 1160  
Db 636 TCCACGAGCAAGATATGGGAGGCGCAG 609

## RESULT 14

US-09-949-016-122372/c  
; Sequence 122372, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 122372  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-122372

Query Match 3.3%; Score 77; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 1.2e-26;  
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 425 AGAAGCCAGCAATATCGAAGTGAAGCAAGAAAGCAAAATTCAGTCCCTTCCCAATGAAT 484  
Db 268 AGAAGCCAGCAATATCGAAGTGAAGCAAGCAAGAAAGCAAAATTCAGTCCCTTCCCAATGAAT 209  
Qy 485 GGTACACGCAAGCAATG 501  
Db 208 GGTACACGCAAGCAATG 192

## RESULT 15

US-09-949-016-122387/c  
; Sequence 122387, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 122387  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human

## US-09-949-016-122387

Query Match 3.1%; Score 73; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 9.8e-25;  
Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 262 TAGAAAACCTAGGAATCTATGATGCTGATGGTGGTGGATTTTGGATGGATGATGCCA 321  
Db 86 TAGAAAACCTAGGAATCTATGATGCTGATGGTGGTGGATTTTGGATGGATGATGCCA 27  
Qy 322 AAGTTTATTAGG 334  
Db 26 AAGTTTATTAGG 14

## RESULT 16

US-09-949-016-236/c  
; Sequence 236, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 236  
; LENGTH: 2208  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-236

Query Match 2.5%; Score 59; DB 4; Length 2208;  
Best Local Similarity 100.0%; Pred. No. 5.2e-18;  
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2239 GGCATCCGGAACCTGACACCAACAGAGAGCGCGCTTCCAGCAATTTAGCATGAATC 2297  
Db 59 GGCATCCGGAACCTGACACCAACAGAGAGCGCGCTTCCAGCAATTTAGCATGAATC 1

## RESULT 17

US-09-949-016-122345/c  
; Sequence 122345, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 122345  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-122345

Query Match 1.9%; Score 45; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 2.8e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 943 CAGAAACAATAGAAAAACAGATGATCCAGAACAAAAAGCAAAAG 987  
|||  
DB 212 CAGAAACAATAGAAAAACAGATGATCCAGAACAAAAAGCAAAAG 168  
|||

RESULT 18  
US-09-949-016-13283  
; Sequence 13283, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13283  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-13283

Query Match 0.9%; Score 22; DB 4; Length 9984;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 CAGCAGCAGCAGCGGCTCCGGC 68  
|||  
DB 920 CAGCAGCAGCAGCGGCTCCGGC 941  
|||

RESULT 19  
US-09-614-034-188  
; Sequence 188, Application US/09614034  
; Patent No. 6489307  
; GENERAL INFORMATION:  
; APPLICANT: PHILLIPS, M. IAN  
; APPLICANT: ZHANG, YUAN  
; TITLE OF INVENTION: ANTISENSE COMPOSITIONS TARGETED TO BETAL-ADRENOCEPTOR-SPECIFIC MR  
; TITLE OF INVENTION: METHODS OF USE  
; FILE REFERENCE: 4300.013900  
; CURRENT APPLICATION NUMBER: US/09/614,034  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 09/152,717  
; PRIOR FILING DATE: 1998-09-14  
; PRIOR APPLICATION NUMBER: PCT/US99/21007  
; PRIOR FILING DATE: 1999-09-14  
; NUMBER OF SEQ ID NOS: 204  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 188  
; LENGTH: 1845  
; TYPE: DNA  
; ORGANISM: UNKNOWN  
; FEATURE:  
; OTHER INFORMATION: SYNTHETIC OLIGONUCLEOTIDE  
; NAME/KEY: misc feature  
; LOCATION: (1828)..(1828)  
; OTHER INFORMATION: WHERE N = A, T, C OR G  
US-09-614-034-188

Query Match 0.9%; Score 21; DB 4; Length 1845;  
Best Local Similarity 100.0%; Pred. No. 9.6;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 CAGCAGCGGCTCCGGCAGCGG 73  
|||  
DB 93 CAGCAGCGGCTCCGGCAGCGG 113  
|||

RESULT 20  
US-09-859-604-10/c  
; Sequence 10, Application US/09859604  
; Patent No. 6835370  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R., Suzanne M  
; APPLICANT: de la Monte, Alan H  
; APPLICANT: Ghanbari, Hossein A  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 CIP  
; CURRENT APPLICATION NUMBER: US/09/859,604  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:Location (-1)  
; OTHER INFORMATION: oligonucleotide  
US-09-859-604-10

Query Match 0.9%; Score 20; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 AATGCCCCAGCGTAAGAATG 30  
|||  
DB 20 AATGCCCCAGCGTAAGAATG 1  
|||

RESULT 21  
US-09-859-604-11/c  
; Sequence 11, Application US/09859604  
; Patent No. 6835370  
; GENERAL INFORMATION:  
; APPLICANT: Wands, Jack R.  
; APPLICANT: de la Monte, Suzanne M  
; APPLICANT: de la Monte, Alan H  
; APPLICANT: Ghanbari, Hossein A  
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS  
; FILE REFERENCE: 21486-032 CIP  
; CURRENT APPLICATION NUMBER: US/09/859,604  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 09/436,184  
; PRIOR FILING DATE: 1999-11-08  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:Location (-6)  
; OTHER INFORMATION: oligonucleotide  
US-09-859-604-11

Query Match 0.9%; Score 20; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 0.9%; Score 20; DB 4; Length 342;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 26  
US-09-949-016-3254  
; Sequence 3254, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3254
; LENGTH: 1106
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3254

Query Match          0.9%; Score 20; DB 4; Length 1106;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      43  GCAACAGCAGCAGCAGCGGC 62
Db      343 GCAACAGCAGCAGCAGCGGC 362
|||||
|||||

RESULT 27
US-09-949-016-1812
; Sequence 1912, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1812
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1812

Query Match          0.9%; Score 20; DB 4; Length 1116;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      43  GCAACAGCAGCAGCAGCGGC 62
Db      353 GCAACAGCAGCAGCAGCGGC 372
|||||
|||||

RESULT 28
US-09-949-016-14996
; Sequence 14996, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14996
; LENGTH: 8841
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14996

Query Match          0.9%; Score 20; DB 4; Length 8841;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14996
; LENGTH: 8841
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14996

Query Match          0.9%; Score 20; DB 4; Length 8841;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      43  GCAACAGCAGCAGCAGCGGC 62
Db      6077 GCAACAGCAGCAGCAGCGGC 6096
|||||
|||||

RESULT 29
US-09-949-016-11878
; Sequence 11878, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11878
; LENGTH: 11668
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-11878

Query Match          0.9%; Score 20; DB 4; Length 11668;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      43  GCAACAGCAGCAGCAGCGGC 62
Db      8971 GCAACAGCAGCAGCAGCGGC 8990
|||||
|||||

RESULT 30
US-09-949-016-13554
; Sequence 13554, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13554
; LENGTH: 11668
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13554

Query Match          0.9%; Score 20; DB 4; Length 11668;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13554  
; LENGTH: 11713  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-13554

Query Match 0.9%; Score 20; DB 4; Length 11713;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCAGCGGC 62  
|||||  
Db 8949 GCAACAGCAGCAGCAGCGGC 8968

## RESULT 31

US-09-949-016-17029  
; Sequence 17029, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17029  
; LENGTH: 47284  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17029

Query Match 0.9%; Score 20; DB 4; Length 47284;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 314 TGATGCCAAAGTTTATTAG 333  
|||||  
Db 20734 TGATGCCAAAGTTTATTAG 20753

## RESULT 32

US-09-949-016-12762/c  
; Sequence 12762, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12762

; LENGTH: 117937  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-12762

Query Match 0.9%; Score 20; DB 4; Length 117937;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 972 GAACAAAAGCAAAAGTTAA 991  
|||||  
Db 50840 GAACAAAAGCAAAAGTTAA 50821

## RESULT 33

US-09-949-016-15775/c  
; Sequence 15775, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15775  
; LENGTH: 117937  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-15775

Query Match 0.9%; Score 20; DB 4; Length 117937;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 972 GAACAAAAGCAAAAGTTAA 991  
|||||  
Db 50840 GAACAAAAGCAAAAGTTAA 50821

## RESULT 34

US-09-513-999C-35281  
; Sequence 35281, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclert, A.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59 US2 REG  
; CURRENT APPLICATION NUMBER: US/09/513,999C  
; CURRENT FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122,487  
; PRIOR FILING DATE: 1999-02-26  
; NUMBER OF SEQ ID NOS: 36681  
; SOFTWARE: Patent.pm  
; SEQ ID NO 35281  
; LENGTH: 148  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-513-999C-35281

Query Match 0.8%; Score 19; DB 4; Length 148;  
Best Local Similarity 100.0%; Pred. No. 88;



Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 AACAGCAGCAGCGGCT 63  
Db 20 AACAGCAGCAGCGGCT 38

## RESULT 35

US-09-248-796A-13462/c  
; Sequence 13462, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstein et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 13462  
; LENGTH: 309  
; TYPE: DNA  
; ORGANISM: Candida albicans  
US-09-248-796A-13462

Query Match 0.8%; Score 19; DB 4; Length 309;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1004 TAACTTTTAAATAAATTT 1022  
Db 168 TAACTTTTAAATAAATTT 150

## RESULT 36

US-09-949-016-61822/c  
; Sequence 61822, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 61822  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-61822

Query Match 0.8%; Score 19; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 47 CAGCAGCAGCAGCGGCTCC 65  
Db 109 CAGCAGCAGCAGCGGCTCC 91

## RESULT 37

US-09-248-796A-4406/c  
; Sequence 4406, Application US/09248796A  
; Patent No. 6747137  
; GENERAL INFORMATION:  
; APPLICANT: Keith Weinstein et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
; FILE REFERENCE: 107196.132  
; CURRENT APPLICATION NUMBER: US/09/248,796A  
; CURRENT FILING DATE: 1999-02-12  
; PRIOR APPLICATION NUMBER: US 60/074,725  
; PRIOR FILING DATE: 1998-02-13  
; PRIOR APPLICATION NUMBER: US 60/096,409  
; PRIOR FILING DATE: 1998-08-13  
; NUMBER OF SEQ ID NOS: 28208  
; SEQ ID NO 4406  
; LENGTH: 951  
; TYPE: DNA  
; ORGANISM: Candida albicans  
US-09-248-796A-4406

Query Match 0.8%; Score 19; DB 4; Length 951;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1004 TAACTTTTAAATAAATTT 1022  
Db 810 TAACTTTTAAATAAATTT 792

## RESULT 38

US-09-949-016-1792/c  
; Sequence 1792, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1792  
; LENGTH: 1599  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-1792

Query Match 0.8%; Score 19; DB 4; Length 1599;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 47 CAGCAGCAGCAGCGGCTCC 65  
Db 963 CAGCAGCAGCAGCGGCTCC 945

## RESULT 39

US-08-339-214-25/c  
; Sequence 25, Application US/08339214  
; Patent No. 6348334  
; GENERAL INFORMATION:  
; APPLICANT: Nagata, Shigikazu  
; APPLICANT: Suda, Takashi  
; APPLICANT: Takahashi, Tomoniro  
; APPLICANT: Nakamura, No. 634833410

;; TITLE OF INVENTION: A Fas Ligand, A Fragment Thereof and DNA  
;; TITLE OF INVENTION: Encoding the Same  
;; NUMBER OF SEQUENCES: 103  
;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Birch, Stewart, Kolasch & Birch  
;; STREET: P.O. Box 747  
;; CITY: Falls Church  
;; STATE: Virginia  
;; COUNTRY: USA  
;; ZIP: 22040-0747

## COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/339,214  
;; FILING DATE: 10-NOV-1994  
;; CLASSIFICATION: 435

;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Murphy Jr, Gerald M.  
;; REGISTRATION NUMBER: 28,977  
;; REFERENCE/DOCKET NUMBER: 1110-139P  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 703-205-8000  
;; TELEFAX: 703-205-8050

;; INFORMATION FOR SEQ ID NO: 25:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 1623 base pairs

;; TYPE: nucleic acid

;; STRANDEDNESS: not relevant

;; TOPOLOGY: not relevant

;; MOLECULE TYPE: cDNA to mRNA

;; HYPOTHETICAL: NO

;; ANTI-SENSE: NO

;; ORIGINAL SOURCE:

;; ORGANISM: rat

;; FEATURE:

;; NAME/KEY: CDS

;; LOCATION: 74...907

;; OTHER INFORMATION: /label= Figs\_2-3

US-08-339-214-25

Query Match 0.8%; Score 19; DB 3; Length 1623;  
Best Local Similarity 100.08; Pred. No. 88;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 812 ACCTCTAGAAATGAAGG 830

Db 1052 ACCTCTAGAAATGAAGG 1034

## RESULT 40

US-09-149-476-110/c

;; Sequence 110, Application US/09149476

;; Patent No. 6420526

;; GENERAL INFORMATION:

;; APPLICANT: Rosen et al.

;; TITLE OF INVENTION: 186 Human Secreted proteins

;; FILE REFERENCE: PZ002P1

;; CURRENT APPLICATION NUMBER: US/09/149,476

;; CURRENT FILING DATE: 1998-09-08

;; EARLIER APPLICATION NUMBER: PCT/US98/04493

;; EARLIER FILING DATE: 1998-03-06

;; EARLIER APPLICATION NUMBER: 60/040,162

;; EARLIER FILING DATE: 1997-03-07

;; EARLIER APPLICATION NUMBER: 60/040,333

;; EARLIER FILING DATE: 1997-03-07

;; EARLIER APPLICATION NUMBER: 60/038,621

;; EARLIER FILING DATE: 1997-03-07

;; EARLIER APPLICATION NUMBER: 60/040,626

;; EARLIER FILING DATE: 1997-03-07

;; EARLIER APPLICATION NUMBER: 60/040,334

;; EARLIER FILING DATE: 1997-03-07  
;; EARLIER APPLICATION NUMBER: 60/040,336  
;; EARLIER FILING DATE: 1997-03-07  
;; EARLIER APPLICATION NUMBER: 60/040,163  
;; EARLIER FILING DATE: 1997-03-07  
;; EARLIER APPLICATION NUMBER: 60/047,600  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,615  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,597  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,502  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,633  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,583  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,617  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,618  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,503  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,592  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,581  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,584  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,500  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,587  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,492  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,598  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,613  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,582  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,596  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,612  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,632  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/047,601  
;; EARLIER FILING DATE: 1997-05-23  
;; EARLIER APPLICATION NUMBER: 60/043,580  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,568  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,314  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,569  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,311  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,671  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,674  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,669  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,312  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,313  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,672  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 60/043,315  
;; EARLIER FILING DATE: 1997-04-11  
;; EARLIER APPLICATION NUMBER: 1997-04-11

EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/056,886  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,877  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,889  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,893  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,630  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,878  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,662  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,872  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,882  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,637  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,903  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,888  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,879  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,880  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,894  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,911  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,636  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,874  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,910  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,864  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,631  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,845  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,892  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,761  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/047,595  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,599  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,588  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,585  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,586  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,590  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,594  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,589  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,593  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,614  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,578  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,576

EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/047,501  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,670  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/056,632  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,664  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,876  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,881  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,909  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,875  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,862  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057,650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056,884  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/057,669  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/049,610  
EARLIER FILING DATE: 1997-06-13  
EARLIER APPLICATION NUMBER: 60/061,060  
EARLIER FILING DATE: 1997-10-02

Query Match 0.8%; Score 19; DB 3; Length 1751;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1836 AAAGTCTCTTCTGCGCTG 1854  
Db 692 AAAGTCTCTTCTGCGCTG 674

RESULT 41  
US-09-949-016-5583  
Sequence 5583, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTIER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5583  
LENGTH: 1821  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-5583

Query Match 0.8%; Score 19; DB 4; Length 1821;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy      36  AGCAGCGGCACAGCAGCA 54
      |||||||
Db      437  AGCAGCGGCACAGCAGCA 455

RESULT 42
US-09-620-312D-49/c
; Sequence 49, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radjoe T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/486,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 49
; LENGTH: 2286
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (69)..(1727)
US-09-620-312D-49
Query Match      0.8%; Score 19; DB 4; Length 2286;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1836  AAAGTCTCTTCCTGCTG 1854
      |||||||
Db      1839  AAAGTCTCTTCCTGCTG 1821

RESULT 43
US-09-149-476-264/c
; Sequence 264, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002F1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/043,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313

```

; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/043,672  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/043,315  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/048,974  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/056,886  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,877  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,889  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,893  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,630  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,878  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,662  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,872  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,882  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,637  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,903  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,888  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,879  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,880  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,894  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,911  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,636  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,874  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,910  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,864  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,631  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,845  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,892  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/057,761  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/047,595  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,599  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,588  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,585  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,586  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,590  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,594  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,589  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/047,593  
 ; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,614  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/043,578  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/043,576  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/047,501  
 ; EARLIER FILING DATE: 1997-05-23  
 ; EARLIER APPLICATION NUMBER: 60/043,670  
 ; EARLIER FILING DATE: 1997-04-11  
 ; EARLIER APPLICATION NUMBER: 60/056,632  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,664  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,876  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,881  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,909  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,875  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,862  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,887  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/056,908  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/048,964  
 ; EARLIER FILING DATE: 1997-06-06  
 ; EARLIER APPLICATION NUMBER: 60/057,650  
 ; EARLIER FILING DATE: 1997-09-05  
 ; EARLIER APPLICATION NUMBER: 60/056,884  
 ; EARLIER FILING DATE: 1997-08-22  
 ; EARLIER APPLICATION NUMBER: 60/057,669  
 ; EARLIER FILING DATE: 1997-09-05  
 ; EARLIER APPLICATION NUMBER: 60/049,610  
 ; EARLIER FILING DATE: 1997-06-13  
 ; EARLIER APPLICATION NUMBER: 60/061,060  
 ; EARLIER FILING DATE: 1997-10-02  
 ;  
 Query Match 0.8% Score 19; DB 3; Length 2377;  
 Best Local Similarity 100.0% Pred. No. 88;  
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1836 AAAGTCTCTCTCTGCGCTG 1854  
 Db 1676 AAAGTCTCTCTCTGCGCTG 1658  
 ;  
 RESULT 44  
 US-09-949-016-2642  
 ; Sequence 2642, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; PRIOR FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2642  
 ; LENGTH: 3462  
 ; TYPE: DNA  
 ; ORGANISM: Human

US-09-949-016-2642

Query Match 0.8%; Score 19; DB 4; Length 3462;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1777	CTTTAGAAAGAAACTGGAA	1795
Dβ	2966	CTTTAGAAAGAAACTGGAA	2984

```

RESULT 45
US-09-949-016-633
; Sequence 633, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 633
; LENGTH: 3482
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(3482)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-633

```

Query Match 0.8%; Score 19; DB 4; Length 3482;  
Best Local Similarity 100.0%; Pred. No. 88;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1777 CTTTAGAAAGAAACTGGAA 1795  
|||  
Db 2975 CTTTAGAAAGAAACTGGAA 2993

```

RESULT 46
US-09-949-016-13534/c
; Sequence 13534, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13534
; LENGTH: 6625
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13534

```

```

Query Match      0.8%; Score 19; DB 4; Length 6625;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 47 CAGCAGCAGCAGCGGCTCC 65
      |||||
Db 3989 CAGCAGCAGCAGCGGCTCC 3971

```

RESULT 47  
US-09-949-016-12375  
; Sequence 12375, Application US/09949016  
; Patent NO. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: C1001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12375  
; LENGTH: 7460  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-12375

	Query Match	0.8%;	Score 19;	DB 4;	Length 7460;
	Best Local Similarity	100.0%;	Pred. No. 87;		
	Matches 19; Conservative	0;	Mismatches	0;	Gaps 0;
Qy	1777 CTTTAGAAGAAACTGGAA	1795			
Dd	4966 CTTTAGAAGAAACTGGAA	4984			

RESULT 48  
US-09-949-016-14384  
; Sequence 14384, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14384  
; LENGTH: 7462  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-14384

```
Query Match          0.8%; Score 19; DB 4; Length 7462;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Caps 0;
```

US-09-949-016-13534

QY 1777 CTTTAGAAGAACTGGAA 1795  
Db 4966 CTTTAGAAGAACTGGAA 4984

RESULT 49  
US-09-949-016-15619  
; Sequence 15619, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15619  
; LENGTH: 7700  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-15619

Query Match 0.8%; Score 19; DB 4; Length 7700;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1777 CTTTAGAAGAACTGGAA 1795  
Db 6898 CTTTAGAAGAACTGGAA 6916

RESULT 50  
US-09-949-016-17325  
; Sequence 17325, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17325  
; LENGTH: 28843  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)...(28843)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-17325

Query Match 0.8%; Score 19; DB 4; Length 28843;  
Best Local Similarity 100.0%; Pred. No. 87;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 AGCAGCGGCACAGCAGCA 54

Db 18759 AGCAGCGGCACAGCAGCA 18777

RESULT 51  
US-09-949-016-11868/c  
; Sequence 11868, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11868  
; LENGTH: 300598  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)...(300598)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-11868

Query Match 0.8%; Score 19; DB 4; Length 300598;  
Best Local Similarity 100.0%; Pred. No. 86;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1072 GCGGAAAAATTGAGGAAGC 1090  
Db 249442 GCGGAAAAATTGAGGAAGC 249424

RESULT 52  
US-09-949-016-14588/c  
; Sequence 14588, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14588  
; LENGTH: 302604  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc.feature  
; LOCATION: (1)...(302604)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14588

Query Match 0.8%; Score 19; DB 4; Length 302604;  
Best Local Similarity 100.0%; Pred. No. 86;



Query Match	0.8%; Score 19; DB 4; Length 308362;
Best Local Similarity	100.0%; Pred. No. 86;
Matches	19; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy	1072 GGGGAAAAATTGAGGAGC 1090 
Db	299264 GGGGAAAAATTGAGGAGC 299246 
RESULT 55	
US-09-949-016-12777/c	
; Sequence 12777, Application US/09949016	
; Patent No. 6812339	
; GENERAL INFORMATION:	
; APPLICANT: VENTER, J. Craig et al.	
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED	
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THERE	
; FILE REFERENCE: CL001307	
; CURRENT APPLICATION NUMBER: US/09/949,016	
; CURRENT FILING DATE: 2000-04-14	
; PRIOR APPLICATION NUMBER: 60/241,755	
; PRIOR FILING DATE: 2000-10-20	
; PRIOR APPLICATION NUMBER: 60/237,768	
; PRIOR FILING DATE: 2000-10-03	
; PRIOR APPLICATION NUMBER: 60/231,498	
; PRIOR FILING DATE: 2000-09-08	
; NUMBER OF SEQ ID NOS: 207012	
; SOFTWARE: FastSeq for Windows Version 4.0	
; SEQ ID NO 12777	
; LENGTH: 828152	
; TYPE: DNA	
; ORGANISM: Human	
; FEATURE:	
; NAME/KEY: misc_feature	
; LOCATION: (1)...(828152)	
; OTHER INFORMATION: n = A,T,C or G	
US-09-949-016-12777	
Query Match 0.8%; Score 19; DB 4; Length 828152;	
Best Local Similarity 100.0%; Pred. No. 85;	
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
Qy	280 ATGATGCTGATGGTGATGG 298 
Db	816467 ATGATGCTGATGGTGATGG 816449 
RESULT 56	
US-09-859-604-13	
; Sequence 13, Application US/09859604	
; Patent No. 6835370	
; GENERAL INFORMATION:	
; APPLICANT: Wands, Jack R.	
; APPLICANT: de la Monte, Suzanne M	
; APPLICANT: Deutch, Alan H	
; APPLICANT: Ghanbari, Hossein A	
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF MALIGNANT NEOPLASMS	
; FILE REFERENCE: 21486-032 CIP	
; CURRENT APPLICATION NUMBER: US/09/859,604	
; CURRENT FILING DATE: 2001-05-17	
; PRIOR APPLICATION NUMBER: 09/436,184	
; PRIOR FILING DATE: 1999-11-08	
; NUMBER OF SEQ ID NOS: 13	
; SOFTWARE: PatentIn Ver. 2.1	
; SEQ ID NO 13	
; LENGTH: 22	
; TYPE: DNA	
; ORGANISM: Artificial Sequence	
; FEATURE:	
; OTHER INFORMATION: Description of Artificial Sequence:Sense	
; OTHER INFORMATION: Oligonucleotide	
US-09-859-604-13	

```
Query Match      0.8%; Score 18; DB 4; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      8 TCGAATGCCCGCGGTAA 25
        |||||
Db       5 TCGAATGCCCGCGGTAA 22

RESULT 57
US-09-313-294A-3024
; Sequence 3024, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 3024
; LENGTH: 286
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6476212 700282183H1
; NAME/KEY: unsure
; LOCATION: 87, 108, 146, 200, 239
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-3024

Query Match      0.8%; Score 18; DB 4; Length 286;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      44 CCAACGACGACGACGCG 61
        |||||
Db       181 CCAACGACGACGACGCG 198

RESULT 58
US-08-482-842B-2
; Sequence 2, Application US/08482842B
; Patent No. 5910480
; GENERAL INFORMATION:
; APPLICANT: KOWAN, Ahmet
; APPLICANT: CHASSIN, Dorine
; APPLICANT: BELLET, Dominique
; TITLE OF INVENTION: NEW PROTEIN CALLED EPI/PLACENTIN,
; TITLE OF INVENTION: PROCESS FOR THE PREPARATION OF THIS PROTEIN AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION CONTAINING SUCH, DNA CODING FOR
; TITLE OF INVENTION: SAID PROTEIN
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,842B
; FILING DATE: 07-JUN-1995
```

```
CLASSIFICATION: 514
PRIOR APPLICATION DATA: FR 94-07191
APPLICATION NUMBER: 13-JUN-1994
FILING DATE: 13-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm K.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 017753-058
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 319 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLSCULE TYPE: DNA (genomic)
US-08-482-842B-2

Query Match      0.8%; Score 18; DB 2; Length 319;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      984 AAAGTTAAGAAAGAGAG 1001
        |||||
Db       88 AAAGTTAAGAAAGAGAG 105

RESULT 59
US-09-513-999C-10601
; Sequence 10601, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 10601
; LENGTH: 452
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 440
; OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-10601

Query Match      0.8%; Score 18; DB 4; Length 452;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1495 GCTTCATCCTGAGGCAC 1512
        |||||
Db       200 GCTTCATCCTGAGGCAC 217

RESULT 60
US-09-270-767-10408
; Sequence 10408, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
```

; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 10408  
; LENGTH: 529  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster  
US-09-270-767-10408

Query Match 0.8%; Score 18; DB 4; Length 529;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GCAACAGCAGCAGCAGCG 60  
Db 136 GCAACAGCAGCAGCAGCG 153

RESULT 61  
US-09-918-686-4/c  
; Sequence 4, Application US/09918686  
; Patent No. 6475739  
; GENERAL INFORMATION:  
; APPLICANT: Brunkow, Mary  
; APPLICANT: Prohl, Sean  
; APPLICANT: Paepker, Bryan  
; APPLICANT: Staehling-Hampton, Karen  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING  
; TITLE OF INVENTION: GENOMIC DELETIONS  
; FILE REFERENCE: 240083.515  
; CURRENT APPLICATION NUMBER: US/09/918,686  
; CURRENT FILING DATE: 2001-07-30  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 557  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-918-686-4

Query Match 0.8%; Score 18; DB 4; Length 557;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1830 AAAGCCAAAGTCTCTTC 1847  
Db 265 AAAGCCAAAGTCTCTTC 248

RESULT 62  
US-09-621-976-16896/c  
; Sequence 16896, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 16896  
; LENGTH: 561  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-621-976-16896

Query Match 0.8%; Score 18; DB 4; Length 561;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2100 AAGGAAGGCTGCAAGATT 2117  
Db 349 AAGGAAGGCTGCAAGATT 332

## RESULT 63

US-09-949-016-36982/c  
; Sequence 36982, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 36982  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-36982

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1909 AAGGAAGAGAAATGAAA 1926  
Db 490 AAGGAAGAGAAATGAAA 473

## RESULT 64

US-09-949-016-40425/c  
; Sequence 40425, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 40425  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-40425

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1602 CTGGGGATGCCATGCAG 1619  
Db 414 CTGGGGATGCCATGCAG 397

RESULT 65  
US-09-949-016-42135  
; Sequence 42135, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 42135  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-42135

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 280 ATGATGCTGATGGTGATG 297  
Db 320 ATGATGCTGATGGTGATG 337

RESULT 66  
US-09-949-016-42136  
; Sequence 42136, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 42136  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-42136

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 280 ATGATGCTGATGGTGATG 297  
Db 280 ATGATGCTGATGGTGATG 297

RESULT 67  
US-09-949-016-42137  
; Sequence 42137, Application US/09949016  
; Patent No. 6812339

; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 42137  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-42137

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 280 ATGATGCTGATGGTGATG 297  
Db 222 ATGATGCTGATGGTGATG 239

RESULT 68  
US-09-949-016-42138  
; Sequence 42138, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 42138  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-42138

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 280 ATGATGCTGATGGTGATG 297  
Db 208 ATGATGCTGATGGTGATG 225

RESULT 69  
US-09-949-016-42139  
; Sequence 42139, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 42139  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-42139

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 280 ATGATGCTGATGGTGATG 297  
Db 205 ATGATGCTGATGGTGATG 222

## RESULT 70

US-09-949-016-42140  
; Sequence 42140, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 42140  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-42140

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 280 ATGATGCTGATGGTGATG 297  
Db 186 ATGATGCTGATGGTGATG 203

## RESULT 71

US-09-949-016-53139/c  
; Sequence 53139, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 53139  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-53139

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1772 AAGTCTTTAGAAAGAAA 1789  
Db 214 AAGTCTTTAGAAAGAAA 197

## RESULT 72

US-09-949-016-67303  
; Sequence 67303, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 67303  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-67303

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 48 AGCAGCAGCAGCGGCTCC 65  
Db 41 AGCAGCAGCAGCGGCTCC 58

## RESULT 73

US-09-949-016-69417/c  
; Sequence 69417, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 69417  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-69417

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 609 GAACCTGAAGTATCTCAT 626  
|||||  
Db 280 GAACCTGAAGTATCTCAT 263

## RESULT 74

US-09-949-016-69418/c  
; Sequence 69418, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 69418  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-69418

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 609 GAACCTGAAGTATCTCAT 626  
|||||  
Db 372 GAACCTGAAGTATCTCAT 355

## RESULT 75

US-09-949-016-69419/c  
; Sequence 69419, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 69419  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-69419

Query Match 0.8%; Score 18; DB 4; Length 601;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 609 GAACCTGAAGTATCTCAT 626  
|||||  
Db 463 GAACCTGAAGTATCTCAT 446

Search completed: March 25, 2005, 22:35:18  
Job time : 417 secs

This Page Blank (uspto)